

27.8
27.8

BRITISH JOURNAL
OF
PSYCHOLOGY

EDITED BY
H. J. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

ASSISTED BY
J. H. B. HARRIS

THE BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY
BORIS SEMEONOFF

WITH THE ASSISTANCE OF
C. BURT

AIDED IN THE SELECTION OF PAPERS BY

F. C. BARTLETT	G. C. DREW	T. H. PEAR
E. A. BOTT	H. J. EYSENCK	G. RYLE
R. S. CREED	L. S. HEARNshaw	W. J. H. SPROTT
J. DREVER	M. G. KENDALL	R. H. THOULESS
	O. A. OESER	

ASSISTANT EDITORS:
IAN M. L. HUNTER (Parts 1-2)
R. M. McKENZIE (Parts 3-4)

VOLUME 51, 1960



CAMBRIDGE
AT THE UNIVERSITY PRESS
1960

CONTENTS OF VOL. 51

PART 1. FEBRUARY 1960

	PAGE
Some individual differences in figural after-effects. By PETER McEWEN and ROBERT S. RODGER	1
Visual discrimination of orientation by octopus: mirror images. By N. S. SUTHERLAND	9
A group version of a perceptual maze test. By ALICK ELITHORN, MYFANWY KERR and JOY MOTT	19
Visual size and familiar size: individual differences. By A. C. McKENNEL .	27
Periodicity in reaction time. By P. H. VENABLES	37
Serial order intrusions in immediate memory. By R. CONRAD	45
Stimulus response theory and expectancy. By MICHEL TREISMAN	49
Construction of a psychological scale of weight from fractionation data. By R. G. CHATTERJEA and RAMANATH KUNDU	61
Degree class and attainment in scientific research. By LIAM HUDSON . . .	67
Inhelder and Piaget's <i>The Growth of Logical Thinking</i> : II. A logician's viewpoint. By CHARLES PARSONS	75
Publications recently received	85
Other publications received	96

PART 2. MAY 1960

Problem-solving, past experience and availability of functions. By PER SAUGSTAD and KJELL RAAHEIM	97
Regression to the mean—a confused concept. By A. D. B. CLARKE, ANN M. CLARKE and R. I. BROWN	105
Recognition of ambiguous and unambiguous visual configurations with short and longer exposures. By CRAIG M. MOONEY	119
The optimal perceptual load in a paced auditory inspection task. By E. C. POULTON	127
Anxiety as a function of psychomotor and social behaviour. By H. SYLVIA ANTHONY	141
Neuroticism in children. II. Identification and measurement of a neuroticism factor. By J. G. THORPE	153
Delinquency, maladjustment and unfavourable ecology. By D. H. STOTT .	157
A projector for colour demonstrations. By J. R. TROTTER	171
Publications recently received	173
Other publications received	192

PART 3. AUGUST 1960

	PAGE
Inhibition: A symposium	193
I. Inhibition as a concept in psychology. By G. W. PILKINGTON and PETER McKELLAR	194
II. The investigation of inhibitory phenomena from the standpoint of comparative psychology. By R. G. A. STRETCH	202
III. Effects of drive and effort on inhibition with reinforcement. By DONALD C. KENDRICK	211
IV. Individual differences in inhibitory potential. By H. GWYNNE JONES	220
V. Inhibition—Pavlov, Hull, Eysenck. By R. L. REID	226
Emergence of set on the basis of imaginal situations. By R. NATADZE .	237
The stroboscopic patterns. III. Further experiments and discussion. By J. R. SMYTHIES	247
Imperception of phi: some further determinants. By VERNON HAMILTON .	257
A study of preferences in feminine beauty. By A. H. ILLIFFE	267
Publications recently received	275

PART 4. NOVEMBER 1960

The accuracy of 'reproduced' conversation. By T. H. PEAR	289
Apparent changes in the visual threshold: central or peripheral? By NORMAN F. DIXON	297
The effects of stimulant and depressant drugs on the Bidwell phenomenon By SATORU AIBA	311
Extraversion, reminiscence and satiation effects. By R. LYNN	319
Association between muscular tension and work output. By D. W. FORREST	325
The effect of distraction on time gap in the estimation of short duration. By R. G. CHATTERJEA	335
The effect of illumination conditions on the effectiveness of light-onset as a reinforcer: a test of the Bevan-Adamson reinforcement theory. By HARRY M. B. HURWITZ	341
A preliminary note on the Brook reaction test. By D. F. BROOK and A. W. HEIM	347
A result in visual aesthetics. By CHRISTOPHER ALEXANDER	357
Publications recently received	373

INDEX OF AUTHORS

	PAGE		PAGE
AIBA, S.	311	McEWEN, P.	1
ALEXANDER, C.	357	McKELLAR, P.	194
ANTHONY, H. S.	141	McKENNELL, A. C.	27
BROOK, D. F.	347	MOONEY, C. M.	119
BROWN, R. I.	105	MOTT, J.	19
CHATTERJEA, R. G.	61, 335	NATADZE, R.	237
CLARKE, A. D. B.	105	PARSONS, C.	75
CLARKE, A. M.	105	PEAR, T. H.	289
CONRAD, R.	45	PILKINGTON, G. W.	194
DIXON, N. F.	297	POULTON, E. C.	127
ELITHORN, A.	19	RAAHEIM, K.	97
FORREST, D. W.	325	REID, R. L.	226
HAMILTON, V.	257	RODGER, R. S.	1
HEIM, A. W.	347	SAUGSTAD, P.	97
HUDSON, L.	67	SMYTHIES, J. R.	247
HURWITZ, H. M. B.	341	STOTT, D. H.	157
ILLIFFE, A. H.	267	STRETCH, R. G. A.	202
JONES, H. GWYNNE	220	SUTHERLAND, N. S.	9
KENDRICK, D.C.	211	THORPE, J. G.	153
KERR, M.	19	TREISMAN, M.	49
KUNDU, R.	61	TROTTER, J. R.	171
LYNN, R.	319	VENABLES, P. H.	37

INDEX OF PUBLICATIONS REVIEWED

- ABERCROMBIE, M. L. J. *The Anatomy of Judgment* 280
- ALEXANDER, F. *The Western Mind in Transition* 375
- ALLPORT, G. W. *Becoming. Basic Considerations for a Psychology of Personality* 384
- AMES, L. B., METRAUX, R. W. & WALKER, R. N. *Adolescent Rorschach Responses* 176
- ANDRY, R. G. *Delinquency and Parental Pathology* 378
- L'Année Psychologique. 1958; 1959* 373
- BANTON, M. *White and Coloured* 180
- BECK, S. J. & MOLISH, H. B. *Reflexes to Intelligence: A Reader in Clinical Psychology* 175
- BIEDMA, C. J. & D'ALFONSO, P. G. *Die Sprache der Zeichnung* 381
- BINDER, H. *Die Helldunkeldeutungen im Psychodiagnostischen Experiment von Rorschach.* 177
- BINDRA, D. *Motivation: A Systematic Reinterpretation* 89
- BÖHM, E. *A Textbook in Rorschach Test Diagnosis for Psychologists, Physicians and Teachers* 187
- BUROS, O. K. (ed.). *The Fifth Mental Measurements Yearbook* 186
- CHERTOK, L. *L'Hypnose: Problèmes Théoriques et pratiques* 93
- CLOSTERMANN, G. *Studien zur Testwissenschaft* 177
- CRONBACH, L. J. *Essentials of Psychological Testing* (2nd edition) 380
- DESCARTES. *Discourse on Method and Other Writings* 384
- DIAMOND, S. *Information and Error* 178
- DIENES, Z. P. *Concept Formation and Personality* 175
- DOLLARD, J. & AULD, F. *Scoring Human Motives: A Manual* 381
- DRAKE, L. E. & OETTING, E. R. *An MMPI Codebook for Counselors* 178
- DULANY, D. E. and others. *Contributions to Modern Psychology: Selected Readings in General Psychology* 92
- EDWARDS, A. L. *The Social Desirability Variable in Personality Assessment and Research* 279
- EYSENCK, H. J. *Manual of the Maudsley Personality Inventory* 185
- FERGUSON, G. A. *Statistical Analysis in Psychology and Education* 286
- GAGNE, R. M. & FLEISHMAN, E. A. *Psychology and Human Performance* 190

- GINSBERG, E. and others. *The Ineffective Soldier: Lessons for Management and the Nation* 284
- GOMBRICH, E. H. *Art and Illusion* 279
- Grundlagenstudien aus Kybernetik und Geisteswissenschaft* 288
- GUILLIAN, G. J. M. *Charcot—1825–1893: His Life—His Work* 173
- HALMOS, P. & ILIFFE, A. *Readings in General Psychology* 92
- HESS, R. D. & HANDEL, G. *Family Worlds* 174
- HIRST, R. J. *The Problems of Perception* 191
- Human Factors: The Journal of the Human Factors Society of America* 287
- JANIS, I. L. and others. *Personality and Persuasibility* 277
- JONES, E. *Free Associations: Memories of a Psychoanalyst* 173
- Journal of Child Psychology & Psychiatry and Allied Disciplines* 287
- JUNG, C. G. *Psychology and Religion* 384
- KHERUMIAN, R. & PICKFORD, R. W. *Hérédité et Fréquence des Dyschromatopsies* 188
- KLEIN, M. *Our Adult World and its Roots in Infancy* 384
- KOCH, S. *Psychology: A Study of a Science*. Study I. Vol. 1. *Sensory, Perceptual and Physiological formulations* 85
- KOCH, S. *Psychology: A Study of a Science*. Study I. Vol. 2. *General Systematic Formulations, Learning and Special Processes* 88
- KOCH, S. *Psychology: A Study of a Science*. Study I. Vol. 3. *Formulations of the Person and the Social Context* 373
- KÖHLER, W. *Dynamische Zusammenhänge in der Psychologie* 94
- KÖNIG, K. *Der Mongolismus* 379
- LEDWITH, N. H. *Rorschach Responses of Elementary School Children* 176
- LEWIS, D. *Quantitative Methods in Psychology* 285
- MARCUSE, F. L. *Hypnosis: Fact and Fiction* 93
- Mechanisation of Thought Processes (Symposium)* 376
- MILLER, D. R. & SWANSON, G. E. *The Changing American Parent* 93
- MILLER, G. A., GALANTER, E. & PRIBRAM, K. H. *Plans and the Structure of Behaviour* 275
- OPIE, I. & P. *The Lore and Language of School-children* 382
- OPLER, M. K. (ed.). *Culture and Mental Health* 377
- OSBORN, R. *Humanism and Moral Theory* 278
- PASAMANICK, B. (ed.) *Epidemiology of Mental Disorder* 282
- PIAGET, J. & INHELDER, B. *La Genèse des Structures Logiques Élémentaires: Classifications et Sériations* 181
- Psychopharmacologia* 94
- REVERS, W. J. *Der Thematische Apperzeptionstest (TAT)* 94
- RIPER, C. V. & IRWIN, J. V. *Voice and Articulation* 283
- ROE, A. & SIMPSON, G. G. *Behaviour and Evolution* 91
- SELLITZ, C. and others. *Research Methods in Social Relations* 380
- SLONIEWSKA, H. *Psychological Aspects of Interest* 383
- SMITH, F. V. *Explanation of Human Behaviour* 277
- STEPHEN, K. *The Wish to Fall Ill* 384
- STOGDILL, R. M. *Individual Behaviour and Group Achievement* 179
- SUTTIE, I. D. *The Origins of Love and Hate* 384
- SZONDI, L., MOSER, V. & WEBB, M. W. *The Szondi Test in Diagnosis* 187
- TANSLEY, A. E. & GULLIFORD, R. *The Education of Slow-learning Children* 378
- TAYLOR, E. M. *Psychological Appraisal of Children with Cerebral Defects* 281
- TERMAN, L. & OGDEN, M. H. *The Gifted Group at Mid-Life* 184
- THIBAUT, J. W. & KELLEY, H. H. *The Social Psychology of Groups* 173
- THORNDIKE, R. L. & HAGEN, E. *10,000 Careers* 382
- TUSSING, L. *Psychology for Better Living* 285
- WATSON, R. I. *Psychology of the Child* 93
- WESTROPP, C. & WILLIAMS, M. *Health and Happiness in Old Age* 285
- ZUCKER, L. J. *Ego Structure in Paranoid Schizophrenia* 190
- ZULLIGER, H. *The Behn-Rorschach Test* 283

SOME INDIVIDUAL DIFFERENCES IN FIGURAL AFTER-EFFECTS

BY PETER McEWEN AND ROBERT S. RODGER

Psychology Department, The Queen's University, Belfast

Three different hypotheses have been proposed to account for individual differences in figural after-effects (Klein & Krech, 1952; Wertheimer & Wertheimer, 1954; Eysenck, 1955). Each implies the existence of a positive correlation between the measures of after-effect in different modalities. The study reported here explored this problem for vision and kinaesthesia. Thirty-two subjects were used and no significant correlation was found ($\alpha = 0.05$). Men showed significantly larger visual figural after-effects than women, a finding analogous to Wertheimer's (1954a) in kinaesthesia. There was no significant rank correlation between either effect and extraversion-introversion as measured by Heron's (1956) sociability index.

I. INTRODUCTION

Many reports on figural after-effects have referred in passing to observations of large individual differences in the size of after-effects. The possibility that the magnitude of effects in one modality might be correlated with that in another was first suggested by Köhler & Dinnerstein (1947) in relation to visual and kinaesthetic after-effects. Confirmation of the existence of such a relationship would clearly have some importance for a theory of figural after-effects. More recently the significance of the problem has been increased by proposals relating the after-effects to what are generally termed personality variables, to metabolic functions and to individual differences in cortical activity. Klein & Krech (1952) propose a construct which they name 'cortical conductivity', which varies with the individual and which contributes to individual differences in cortical integration and in behaviour. They regard kinaesthetic after-effects as an indirect measure of cortical conductivity and, from work by Klein, they infer that patients with brain injuries have a lowered conductivity and will show after-effects more readily, in greater degree and with greater persistence than normals. They compared their brain-injured group with a normal group on kinaesthetic after-effects and found results, on the four measures used, to be consistently in favour of their hypothesis.

Eysenck (1955) identifies figural after-effects with effects of reactive inhibition. Individual differences in the rate at which reactive inhibition is produced and dissipated and in its strength, predispose individuals to the development of extraverted and introverted patterns of conduct and consequently underlie the type of neurotic disorder which appears if breakdown occurs. He used Klein and Krech's apparatus and procedure on two groups of neurotic patients, one hysterics and the other dysthymics. He found that the hysterics developed after-effects more quickly, in greater magnitude and recovered more slowly than the dysthymics. His views on after-effects and reactive inhibition are given more fully in a later book (Eysenck, 1957).

Wertheimer & Wertheimer (1954) argue that individual differences in the size of figural after-effects reflect differences in cortical modifiability and can therefore be used as a measure of an individual's generalized cortical modifiability. Further, both satiation theory (Köhler & Wallach, 1944) and statistical theory (Osgood & Heyer, 1952) involve

physico-chemical alterations in the tissue of the cortex which could be interpreted as implying metabolic changes. From this Wertheimer and Wertheimer argue that high figural after-effects reflect high physico-chemical modifiability. This leads them to coin a term, 'metabolic efficiency', to refer to the normal range of metabolic function. Figural after-effects are said to be maximal in the normal range of metabolic functioning and to fall off above and below this range. They make certain predictions about the relation of the size of an individual's after-effects and physiological indicators of metabolic efficiency, such as reaction time and sensori-motor learning. Finally they also predict low after-effects in schizophrenics whom they consider to show low metabolic efficiency and they predict a correlation between after-effects in different modalities. These predictions appear to be supported, on the whole, by the evidence derived from experimental tests reported by Wertheimer (1955). He compared the magnitude of figural after-effects with reaction time, basal metabolic rate, thyroid function, circulatory structure and somatotype. He found that they were related in a fashion predictable on the assumption that 'metabolic efficiency' was a common determinant. The one prediction that was not confirmed was of a positive relation between modifiability in a simple sensori-motor task and figural after-effects; the results here indicated a negative correlation. In the same paper he reported data showing that schizophrenics, taken to be a group exhibiting low metabolic efficiency, had low visual and kinaesthetic after-effects.

II. THE PROBLEM

The research to be reported here is concerned with the predicted correlation between figural after-effects in different modalities, namely vision and kinaesthesia. It is clear that each of the proposals outlined above implies the existence of such a correlation, though only Wertheimer and Wertheimer make an explicit prediction that a positive correlation will be found. Each of the theories would be embarrassed by a negative finding. Wertheimer (1955) reported evidence of a positive relationship, significant by the Mann-Whitney U-test. Encouraging though this finding is, it cannot be regarded as settling the issue. In their 1947 paper, Köhler & Dinnerstein admit that their original intention had been to investigate this problem of correlation between visual and kinaesthetic after-effects, but that it soon became plain that this could not be done until more was known about the latter. The situation in this respect is not much improved at the present time and the need for caution has been re-emphasized recently by findings reported by Broadbent (1958). Both Klein & Krech (1952) and Eysenck (1955) had subjects make their judgements on a long wedge over which they moved their fingers until they found the point at which the width seemed equal to that of the $1\frac{1}{2}$ -in. standard. Eysenck (and apparently Klein & Krech also) always presented this wedge with the pointed end towards the subject. Broadbent repeated Eysenck's procedure, but with the wedge alternately pointing towards and away from the subject. He found no correlation between extraversion and after-effect when measured in this fashion. He did find a progressive shift of the P.S.E. towards the subject's body, the extent of which was correlated with extraversion. An increasing shift towards S's body from pre-satiation to post-satiation trials could, in Klein & Krech and Eysenck's experiments, be mistaken for a figural after-effect.

Whether Broadbent's work alters the meaning of the previous findings or not, it makes

clear that there are likely to be unknown variables involved in the measurement of figural after-effects by the Klein & Krech method. Köhler & Dinnerstein did not use a continuously tapering scale for their experiment, but had a scale which tapered in discrete steps. Subjects began at the wide and narrow ends in alternate comparisons. Wertheimer, who uses this type of scale, does not say whether this latter procedure was followed. Köhler & Dinnerstein, initially, and Wertheimer used two inspection-blocks, one in each hand in the stimulation period, a procedure which, Köhler & Dinnerstein point out, does not give a true measure of kinaesthetic after-effects. With these considerations in mind, an attempt was made to improve upon previous procedures for measuring the effects.

III. APPARATUS

(i) *The kinaesthetic after-effect*

The inspection and test blocks were the same as those used by Klein & Krech and Eysenck, $2\frac{1}{2}$ and $1\frac{1}{2}$ in. in width respectively. Instead of either comparison scale, we used a new instrument of our own devising. This consisted of a wood-worker's vice with suitable strips of smoothed hardwood bolted to the jaws. A dial gauge was attached which gave readings of the gap between the strips. The apparatus is shown diagrammatically in Fig. 1.

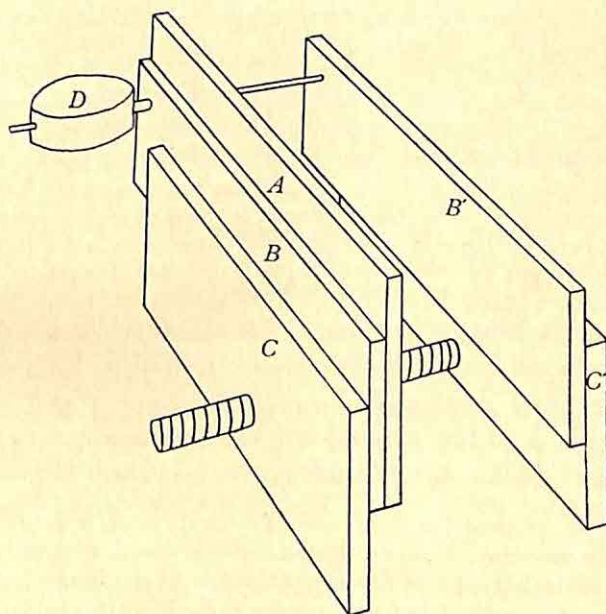


Fig. 1. A, rider mounting; B, B', variable test strips; C, C', jaws of the vice; D, dial gauge.

Riders were used to align the subject's thumb and forefinger on inspection, test and measuring block. The rider itself, not shown in the diagram, was attached on the measuring block over the section marked A. The subject placed his thumb at B and his index finger on the outside of the strip marked B'. The sections marked B' and C' could be moved by means of the worm-screw of the strip marked B'. As these moved, the lever of the dial gauge was displaced by amounts which could be read off the scale.

(ii) *The visual after-effect*

It was desirable to choose inspection and test-figures, distances, sizes and other conditions so as to obtain after-effects as large as possible. Fortunately, for visual after-effects, enough is known about some of these matters for a reasonable choice to be made. In a review of Japanese research, Sagara &

Oyama (1957) have reported that, with a circle as I-figure and two smaller circles, one falling inside the area of the I-circle, as T-figures, maximal effects are obtained when the diameter of the inspection circle is twice that of the affected test circle. Apparatus for this measurement was available from a previous study and has been more fully described in a report (McEwen, 1959). The inspection figure was a ring of light, 61 mm. in inner diameter. The fixation mark, a small red spot of light, was 10 cm. to the right of the centre of the figure as viewed by S. The test figures were two discs of light, varied in size by means of iris diaphragms. The figures were lighted by a concealed lamp behind the screens, the light falling on discs of parchment inserted between the casing and the leaves of the iris diaphragms. The fixation point was between the two discs, 10 cm. from the centre of each. The standard (affected) T-figure was on the subject's left and remained constant at 30.5 mm. diameter. The neutral T-circle was varied by the experimenter and its size could be read off a protractor at the back of the screen on which the diaphragm was mounted. The neutral T-circle (variable) was objectively equal to the standard when the protractor reading was 75° , higher readings indicating smaller circle sizes and vice versa. The T-screen was set directly in front of the head rest in which S's head was placed, the I-screen 30° to the right. The head-rest could be rotated through this angle by the subject. Both screens were 150 cm. from S's eye. During the experiment the room was darkened.

IV. SUBJECTS AND PROCEDURES

The subjects were sixteen men and sixteen women, all but one being undergraduates attending a first course in psychology and naïve as regards figural after-effects. The exception was a research worker in the department who had not taken part in any previous figural after-effect experiments and who was familiar with the problem only in a general way. Her results would have been discarded if they had been in any way extreme or discrepant.

Since it was desirable that measurements of both visual and kinaesthetic after-effects be made in the course of a single session, and since it was conceivable that either practice or stimulation in the first modality might affect performance in the second, half the subjects had the visual experiment first, the other half the kinaesthetic first. There is also evidence, in the case of the kinaesthetic effect, which suggests that the stimulation involved in obtaining control measurements creates some satiation. Therefore, for half the subjects control measurements preceded the post-satiation measurements and for the other half the control measurements were made second, with an interval to allow the experimental satiation effects to subside.

With the aid of a table of random numbers (Fisher & Yates, 1943) subjects, men and women separately, were allocated to one of four orders according to whether they were to have the visual or kinaesthetic task first and control (no inspection figure) or experimental treatment first. The groups thus formed can be labelled VC, VE, KC and KE (representing Visual Control first, Visual Experimental first and so on).

All subjects, as a group were initially given part II of the inventory described by Heron (1956) from which we could derive a sociability (extraversion-introversion) index.

Subjects were given a practice session, in both tasks, lasting about 45 min., before the experiment proper was begun. Practice and experimental sessions were on separate days. Procedures during practice were the same as those used in the subsequent experimental session except that, during the practice session, subjects were given knowledge of results and no inspection figure was used. Points of subjective equality were determined by a constant method and subjects were asked to say whether the variable was larger or smaller than the standard. Judgements of 'equal' were permitted but discouraged and, though such judgements were recorded, they were discarded in calculating the p.s.e. A constant method seemed preferable to an adjustment method because it minimized the time taken between inspection and judgement. The subject continued judging until he had given five judgements of 'greater' and/or 'less' for each preselected value of the variable. In a few cases, there were inversions in the order of judged variate values and we continued to present these stimuli until ten judgements of greater and/or less were given for each.

Three minutes of dark adaptation was allowed prior to the visual task and subjects were blindfolded before encountering the apparatus for the kinaesthetic task. In the latter task, the standard block was presented to the dominant hand (only one of our thirty-two subjects was left-handed).

Immediately prior to the experimental session in each task a further short practice period was given which, at the same time, provided a rough estimate of the subject's p.s.e. Where this practice series was followed by the control treatment, the estimate of the p.s.e. provided a mid-point for the range of variable settings used. When the practice was followed by the experimental treatment, the p.s.e. provided a

starting-point for a descending series, continued until a judgement of smaller was given. From this point onwards, the series was random, as in the control series. In the visual task, the unit used was 2° on the protractor scale representing 0.9 mm., in the diameter of the variable. In the kinaesthetic task, the unit was 0.04 in. In both tasks the initial inspection time was 2 min. and all subsequent inspections lasted 30 sec. On the average, the time from the end of one inspection to the start of the next one was 5 sec. in the visual task and 10 sec. in the kinaesthetic one. In those groups which received the Experimental treatment followed by the Control treatment (the E-C groups), a rest interval of 7 min. was given to allow the effects of satiation to dissipate.

There is another matter which ought to be considered at this stage, namely the adequacy of subjects' fixation during the visual experiment. Eysenck (1957) states that there are difficulties in inducing neurotic patients to fixate the inspection figure properly during the satiation period. He quotes Nicholls (1955) as the basis for, or in support of, this statement. Nicholls, however, did not observe this himself (if he did, he does not say so) but advanced the statement as a *post hoc* hypothesis to account for the discrepancy between his findings and those predicted from Eysenck's theory. Nicholls's suggestion was based on a personal communication from Granger, who used groups of dysthymic and hysteric subjects (as did Nicholls) and observed that hysterical subjects tended to report the disappearance of the fixation point from time to time. Changes in the appearance of the fixation mark and of the inspection figures and partial or complete temporary disappearance has been frequently reported, most fully by Marks (1949). It occurs apparently as a normal consequence of prolonged fixation but there is no observational evidence that any subjects or class of subjects actually move their eyes in such circumstances.

Our subjects were given a practice period before and during which their attention was fully drawn to the likelihood of these phenomenal changes. They were assured that these were usual and encouraged not to move their eyes. They were also told that, at first, they would probably experience eye movements, that with practice these could be avoided but that subjects were to inform the experimenters when eye movements occurred. Such reports decreased rapidly during practice but we continued to ask about eye movements throughout practice and experimental sessions. During the latter, reports were so rare that we did not consider eye movements of which subjects were conscious to be serious sources of error.

V. RESULTS

(i) *Figural after-effects*

The measures of figural after-effects in the visual and kinaesthetic tasks were determined for each subject. The mean effect in the visual test was 9.09° as measured on the protractor scale (5 min. of visual angle approximately) and standard error 1.84° . The mean effect in the kinaesthetic task was 0.166 in. (12% distortion) and standard error 0.034 in. Both figural after-effects are therefore highly significant ($P < 0.001$) and the size of the visual effect is reasonably large as we intended it should be.

Table 1. *Visual figural after-effects data*

Analysis of variance	D.F.	Corrected sum squares	Mean square	F
Between sexes	1	166.53	166.53	8.57**
Between treatments	3	31.59	10.53	—
Interaction S $\times$ T	3	36.85	19.43	—
Residual	24	487.75	—	—
Total	31	722.72	—	—

** Significant at the 1% level.

An hypothesis was set up that there would be no significant differences between the sexes or the treatments in either test and a 5% rejection area was fixed. Two analyses of variance were carried out, one on each set of test results. These are shown in Tables 1 and 2.

The means underlying the significant sex difference in the visual task are men = 11.4,

and women = 6.7. The means underlying the significant treatment difference in the kinaesthetic task are VC = 2.2, KC = 1.8, VE = 0.84 and KE = 0.87.

When these means are tested we find that VC does not differ significantly from KC nor VE from KE, but each of the C pairs differs from each of the E pairs. This difference seems to us to be due to kinaesthetic satiation persisting during the control session in the groups which received the E-C treatment, despite the 7-min. interval, allowed for dissipation. The sex difference in the visual effect is similar to that reported by Wertheimer (1954*b*). Further investigation of this might be worthwhile particularly as Wertheimer (1954*a*) reports a similar difference in kinaesthesia. Although the sex difference in our kinaesthetic task is in the same direction as that reported by Wertheimer, it is not significant.

Table 2. *Kinaesthetic figural after-effects data*

Analysis of variance	D.F.	Corrected sum squares	Mean square	F
Between sexes	1	0.300	0.300	—
Between treatments	3	11.263	3.754	4.34*
Interaction S × T	3	4.206	0.864	—
Residual	24	19.127		
Total	31	34.896	—	—

* Significant at the 5% level.

(ii) *Bilateral kinaesthetic and half-meridional differences*

Wertheimer (1954*a*) found that a width held between thumb and forefinger of the dominant hand was judged smaller than an equal width held in the other hand. He called this a bilateral kinaesthetic difference (BKD) and found a sex difference in the magnitude of this effect. It is also known that there are differences in judged sizes between the two halves of the visual field which Brown (1953) called the half-meridional difference (HMD) though no sex differences were reported. In view of these findings, two further analyses of variance were carried out on our control readings. These are shown in Tables 3 and 4.

Table 3. *Visual control data (P.S.E.)*

Analysis of variance	D.F.	Corrected sum squares	Mean square	F
Between sexes	1	66.12	66.12	4.08
Between treatments	3	206.75	68.92	4.26*
Interaction S × T	3	61.13	16.19	—
Residual	24	376.00		
Total	31	710.00	—	—

* Significant at the 5% level.

Table 4. *Kinaesthetic control data (P.S.E.)*

Analysis of variance	D.F.	Corrected sum squares	Mean squares	F
Between sexes	1	1.09	1.09	—
Between treatments	3	8.69	2.90	—
Interaction S × T	3	10.40	3.05	—
Residual	24	71.91		
Total	31	92.09	—	—

The kinaesthetic data do not show a significant sex difference but the results are in the same direction as that reported by Wertheimer, namely, that women show a greater BKD than men. The sex difference in the visual data is almost significant. Subjects tend to set the variable smaller than the standard, women more markedly than men. The only significant result in Tables 3 and 4 is a visual treatment difference. The appropriate means and standard error are $VC = 74.875$, $KC = 76.500$, $VE = 77.875$, $KE = 81.750^\circ$ and standard error = 2.01° . In view of the fact that kinaesthetic satiation persisted longer than we expected, we may be justified in combining the results of VC with KC (which do not differ significantly) and VE with KE (which do not differ significantly). When this is done the resultant joint means are $C = 75.69$ and $E = 79.81^\circ$ with standard error of 1.41° , producing a difference significant at the 5% level. It seems that visual satiation also persisted longer than expected. Though this did not produce significant treatment differences in Table 1, the results are in the appropriate direction.

(iii) Correlation

Since significant treatment differences have been found, it is not advisable to combine all the data into one bivariate distribution. We should remove the treatment difference in the kinaesthetic measures before correlating them with the visual measures (this is done by writing each measurement as a deviation from its group mean instead of the general mean). When this is done, the product-moment correlation between kinaesthetic and visual effects is $r = +0.12$. If we remove the sex difference in the visual effect, this correlation becomes $r = 0.03$. For $N = 32$, neither of these coefficients differ significantly from zero. It might be of interest to report the coefficients of men and women separately. With 'purified' kinaesthetic measures these are $r_{\text{men}} = +0.04$ $r_{\text{women}} = +0.11$ ($N = 16$ for each). This apparent lack of a correlation between the magnitudes of individual differences in kinaesthetic and visual after-effects could be due to an injudicious selection of subjects. From the point of view of one writer, Eysenck (1955, 1957), further illumination of this issue is gained from the scores our subjects obtained on the sociability (extraversion-introversion) scale of Heron (1956). The appropriate frequency distributions are given in Table 5: a low score is given to extraverts and a high one to introverts.

Table 5. *Distribution of scores on the Heron inventory*

Score ...	1	2	3	4	5	6	7	8	9	10
Men	1	2	1	3	2	4	1	2	0	0
Women	0	1	5	1	4	1	0	1	1	2
Total	1	3	6	4	6	5	1	3	1	2

If this absence of correlation between figural after-effects in these two modalities is confirmed by subsequent research, some alternative explanation may still relate the magnitude of figural after-effects in one or some modalities to introversion-extraversion even though this may not be possible for all modalities.

The relationship between introversion-extraversion, as measured by Heron's scale and the magnitude of figural after-effects in vision and kinaesthesia separately can be measured by computing the coefficient τ (Kendall, 1948). We ranked our data so that large figural after-effects and low questionnaire scores were both given for low rank and used the

formula for tied ranks. Since those subjects who were given the E-C treatment had reduced kinaesthetic figural after-effects, it seemed wise to find Z-scores for the E-C group and the C-E group separately, rank all resulting Z values and correlate these ranks with Heron ranks. The resulting coefficient is $\tau = -0.02$. The rank coefficient of Heron scores and the size of visual effect is $\tau = -0.02$, and when the sex difference in visual effect is removed by using sub-group Z-scores this becomes $\tau = 0.00$. All these coefficients are based on $N = 32$ cases and are obviously non-significant. These calculations provide no support for an hypothesis of relationship between extraversion and figural after-effect in the two modalities studied.

REFERENCES

- BROADBENT, D. E. (1958). Extraversion and the kinaesthetic figural after-effect (Abstract). *Bull. Brit. Psychol. Soc.* No. 35.
- BROWN, K. T. (1953). Methodology of studying figural after-effects and practice effects in the Müller-Lyer illusion. *Amer. J. Psychol.* 66, 629-34.
- EYSENCK, H. J. (1955). Cortical inhibition, figural after-effect and theory of personality. *J. Abnorm. Soc. Psychol.* 51, 94-106.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.
- FISHER, R. A. & YATES, F. (1943). *Statistical Tables for Biological, Agricultural and Medical Research*. Edinburgh: Oliver and Boyd.
- HERON, A. (1956). A two-part personality measure for use as a reasearch criterion. *Brit. J. Psychol.* 47, 243-51.
- KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Griffin.
- KLEIN, G. S. & KRECH, D. (1952). Cortical conductivity in the brain-injured. *J. Personality*, 21, 118-48.
- KÖHLER, W. & DINNERSSTEIN, D. (1947). Figural after-effects in kinaesthesia. In *Miscellanea Psychologica A. Michotte*. Louvain, pp. 196-220.
- KÖHLER, W. & WALLACH, H. (1944). Figural after-effects: an investigation of visual processes. *Proc. Amer. Phil. Soc.* 88, 269-357.
- MCEWEN, P. (1959). Figural after-effect; retinal size and apparent size. *Brit. J. Psychol.* 50, 41-7.
- MARKS, M. R. (1949). Some phenomena attendant on long fixation. *Amer. J. Psychol.* 62, 382-6.
- NICHOLLS, E. G. (1955). The relation between certain personality variables and the figural after-effect. (unpubl. Ph.D. thesis) Univ. of London.
- OSGOOD, C. E. & HEYER, A. W. (1952). A new interpretation of figural after-effects. *Psychol. Rev.* 59, 98-118.
- SAGARA, M. & OYAMA, T. (1957). Experimental studies on figural after-effects in Japan. *Psychol. Bull.* 54, 327-38.
- WERTHEIMER, M. (1954a). Constant errors in the measurement of figural after-effects. *Amer. J. Psychol.* 67, 543-46.
- WERTHEIMER, M. (1954b). The differential satiability of schizophrenic and normal subjects: a test of a deduction from the theory of figural after-effects. *J. Gen. Psychol.* 51, 291-9.
- WERTHEIMER, M. (1955). Figural after-effect as a measure of metabolic efficiency. *J. Personality*, 24, 56-73.
- WERTHEIMER, M. & WERTHEIMER, N. (1954). A metabolic interpretation of individual differences in figural after-effects. *Psychol. Rev.* 61, 279-80.

(Manuscript received 22 September 1958)

VISUAL DISCRIMINATION OF ORIENTATION BY OCTOPUS: MIRROR IMAGES

By N. S. SUTHERLAND

Institute of Experimental Psychology, Oxford University

Octopuses were trained to discriminate visually between different pairs of mirror image shapes. The successive method of discrimination training was used. It was found first that octopuses discriminate more readily between up-down mirror images than between left-right mirror images; secondly, that octopuses discriminate as readily between mirror image shapes as between non-mirror image shapes of equivalent complexity. It was also shown that it is possible to train octopuses on a comparatively difficult discrimination by the method of making reward conditional upon an attack upon the positive shape presented upon its own instead of presenting reward with the positive shape every positive trial. Although initial learning by the former method is slower, a smoother learning curve is obtained by this method, and the final level of performance reached is higher.

I. INTRODUCTION

It has been found (Sutherland, 1957*a*, 1958*b*) that octopuses discriminate very readily between horizontal and vertical rectangles, less readily between horizontal or vertical rectangles and an oblique rectangle, and scarcely at all between oblique rectangles at 90° to one another. It was suggested (Sutherland, 1957*a*) that these findings could be explained in one of two ways: (1) Octopuses might be analysing the horizontal or vertical extents of figures (but not both) relative to their total size. (2) Octopuses might be analysing the extents of figures along both vertical and horizontal co-ordinates. Either hypothesis would account for the findings instanced above. If hypothesis (1) is correct, octopuses should be able to discriminate between figures whose extents differ along the co-ordinate they are analysing, but do not differ along the other co-ordinate; they should be unable to discriminate between the same pair of figures each rotated through 90°, since the figures will now differ only on the co-ordinate along which they are not analysing. If hypothesis (2) is correct then both pairs of figures should be equally discriminable, since provided the difference between the members of each pair is the same it should not matter along which co-ordinate the difference falls. These conditions are satisfied by pairs of figures which are right-left and up-down mirror images of one another. Such a pair of figures differs *either* in their vertical extents (right-left mirror images) *or* in their horizontal extents (up-down mirror images).

The experiments to be reported here constitute a test between the two hypotheses by assessing the discriminability of two sets of mirror image shapes, namely $\top$ $\perp$, and $\vdash$ $\dashv$; $\sqcup$ $\sqcap$ and $\square$ $\square$. If hypothesis (1) is correct, either the two up-down pairs of mirror images should be discriminable but not the two left-right pairs of mirror images or the two left-right pairs of mirror images should be discriminable but not the two up-down pairs of mirror images. If hypothesis (2) is correct, then all pairs of mirror images should be equally discriminable, since on hypothesis (2), the octopus is analysing both horizontal and vertical extents of figures.

The experiments have a wider interest than this since it is sometimes supposed (e.g. Deutsch, 1955) that animals have great difficulty in discriminating mirror images. In

order to test this possibility a further pair of shapes was included in the present experiments which were not mirror images of one another, namely two $\sqcup$ shapes differing from one another not by 180° rotation (mirror images) but by 90° rotation. A further purpose of the present experiment was to investigate the possibility of using more rapid methods of training in discrimination experiments with octopus. A brief preliminary report on part of experiment I ($\top$ shapes) has already appeared (Sutherland, 1957*b*), but it is necessary to report this experiment in more detail here since (a) a false conclusion was drawn from the results in the earlier report; (b) it is desired to compare the different methods used in the two experiments.

II. METHODS

A full account of the experimental situation will be found in Sutherland (1957*a*). In summary, animals were kept in individual tanks, and shapes were presented one at a time (successive discrimination training) at the end of the tank away from the octopus's home. Animals were trained to attack one shape and not to attack the other. Shapes were presented on the end of a transparent Perspex rod, and were moved up and down in the water by hand through a distance of about 3 cm. three times a second. Animals were rewarded for attack on one shape (the positive) by being given a piece of sardine weighing 5 g., and punished for attack on the other shape by an electric shock (10 volts A.C.). Negative shapes remained in the water for 30 sec., positive shapes remained in the water for 30 sec. unless attacked earlier, when they were withdrawn as soon as the animal had been fed after attack. Any attack on the positive shape within 30 sec. was counted as a correct response, any failure to attack the negative shape within 30 sec. was counted as a correct response. In experiment II, shapes remained in the water for 20 sec. only. All animals were trained to attack sardines, and were also given ten trials pre-training to attack the positive shape for a sardine reward before discrimination training proper began. For further details of the experimental procedure, see Sutherland (1957*a*), and of the habits of octopuses, see Boycott (1954).

III. EXPERIMENT I

Method

Six animals (group A) were trained to discriminate between the figures $\top$ and $\perp$. Three animals were trained with $\top$ positive, three with $\perp$ positive. Seven animals (group B) were trained to discriminate between $\neg$ and $\vdash$; four were trained with $\neg$ positive, three with $\vdash$ positive. The cross stroke of the $\top$ shapes used was 10×2 cm. and the tail was 4×2 cm. Ten training trials were given a day, and these always consisted of five positive and five negative trials. For the first 8 days of training a pair of shapes (one positive, one negative) were presented at 2-hr. intervals throughout the day. The presentation of the two members of each pair was separated by at least 10 min. The shape which was presented first within each pair varied from pair to pair in a random manner. During these 8 days' training, octopuses which failed to attack the positive shape within 30 sec. were presented with the piece of sardine with the shape still in the water. They would then attack both shape and sardine together, and after they had seized the sardine, the shape was removed. During a further 4 days of training, the sardine was not presented in association with the shape if the octopus failed to attack within 30 sec.

Results

Learning curves for groups A and B are shown in Fig. 1. It will be seen that though some learning did occur, discrimination did not reach a high level for either group. Moreover, for group A all improvement in performance occurred during the first twenty trials, and group B only improved slightly after the first twenty trials. Nevertheless discrimination over the first 8 days was significantly better than chance for both groups: χ^2 values for the difference between the ratios of attacks on positive shape to total trials and attacks on negative shape to total trials were 11.0 for group A (D.F. = 1, $P < 0.01$), 6.6 for group B (D.F. = 1, $P < 0.02$). There was no significant difference between the two groups over the

first eight days of training ($t = 1.0$, D.F. = 11, $P < 0.3$). Nor were differences between groups trained positive on different members of each pair of shapes significant. Thus in group A, octopuses trained positively on the upright Γ shape averaged 60% correct responses, whereas octopuses trained positively on the inverted Γ shape averaged 57%; the corresponding figures for group B were 57% and 54% correct responses.

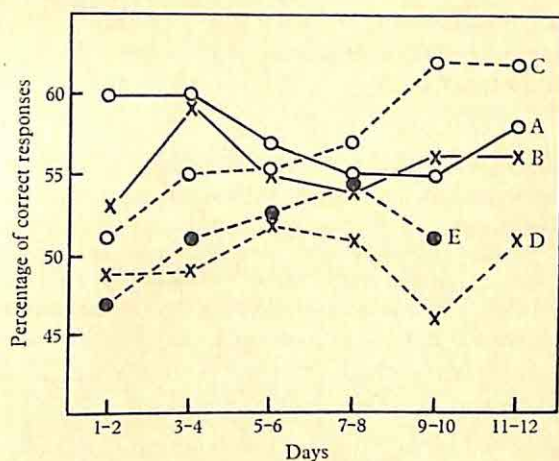


Fig. 1. Initial learning. — groups A, B; Γ figures. - - - groups C, D, E; $\sqcup$ figures. ○ groups A, C; up-down mirror images. × groups B, D; right-left mirror images. ● group E. 50% represents the chance level of response.

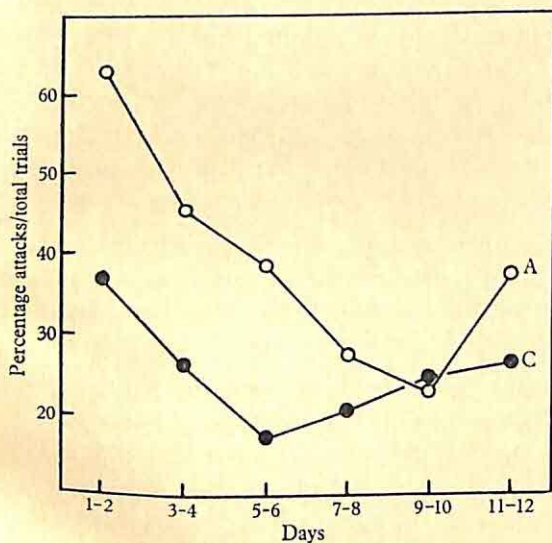


Fig. 2

Fig. 2. The percentage of trials on which attacks were made.

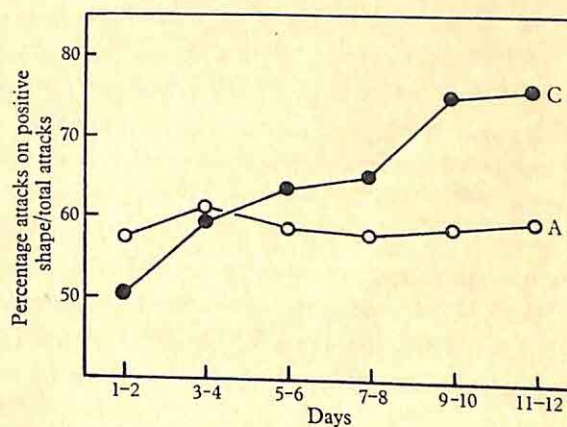


Fig. 3

Fig. 3. The percentage of all attacks made on the positive shape. N.B. 50% represents the chance level of response.

One striking feature of the results of the first 8 days' training is the decline in the total number of attacks made per day (summing attacks made on positive and negative shapes) (see Fig. 2). This decline in the total number of attacks is in fact responsible for the

decline in frequency of correct responding. If an animal attacks very seldom the ratio of correct to total responses will be less than for an animal attacking 50% of the time, even though the ratio number of attacks on positive figure to number of attacks on negative figure is the same. This is clearly revealed by comparison of the curves for performance of group A over the first 8 days of training shown in Figs. 1, 2 and 3. (Figs. 2 and 3 only show curves for group A, since this group will subsequently be compared with group C, but the curves for group B are very similar though slightly displaced downwards.) Fig. 1 shows that the percentage of correct to total responses tends to decline, Fig. 2 shows the marked decline in the total number of attacks made, and Fig. 3 shows that the ratio attacks on positive figure to total attacks remains approximately constant. This last method of plotting scores is perhaps a more accurate measure of discrimination than the first since it omits effects due to the general level of attack decreasing, and reflects only the difference in number of attacks on the positive and on the negative figure. The results raise a further question, why performance fails to increase after the first twenty trials even when it is plotted as in Fig. 3 as the percentage of attacks on the positive of total attacks made. The answer may be that octopuses are learning to wait until the reward is inserted with the shape. Animals which wait until the expiry of 30 sec. will be reinforced for not attacking before the sardine is presented on 50% of trials; animals which attack before 30 sec. will be reinforced on 60% of trials, punished on 40% (assuming that 60% of all attacks are made on the positive figure as they in fact were). If some animals are learning to solve the problem after the first few trials by waiting to attack until the sardine is presented, then this explains why discrimination fails to improve.

As a test for these assumptions a new method of training was adopted from day 9-12: sardine reward was only given if an octopus attacked the positive shape before 30 sec. Fig. 2 shows that this procedure did lead to a slight rise in the total number of attacks, and a corresponding rise in the percentage of correct responses to total trials occurred (Fig. 1). However, Fig. 3 shows that there was no improvement in the discriminatory performance as measured by the ratio attacks on positive to total attacks.

In view of the failure of discriminatory performance to rise above the level attained in the first twenty trials, it is unsafe to conclude on the basis of this experiment alone that there is no difference in ease of learning of up-down and right-left mirror images: if animals were learning not to attack until the expiry of 30 sec., this would tend to conceal any effects due to possible differences in discriminability of the shapes. Experiment II was therefore conducted to determine whether any difference in discriminability between up-down and right-left mirror images does exist and also to determine the practicability of using a new training method in visual discrimination by octopus with shapes of low discriminability.

IV. EXPERIMENT II

Method

Three groups of eight animals each were trained to discriminate between  figures. The cross stroke of the  figures was 7 x 2 cm. and the two tail strokes were each 4 x 1.5 cm. Group C was trained for 12 days to discriminate between an upright and an inverted . Group D was trained for 12 days to discriminate between  and . Group E was trained to discriminate between  and ,  and ,  and  (two animals were trained on each pair). Unfortunately it was only possible to complete 10 days' training with this group. Within each group half the animals were trained with one shape positive,

half with the other shape positive. Six animals (three from group A and three from group B) were given an additional 18 days' training with the same shapes as were used during the first 12 days' training.

All animals were given twenty trials a day. Trials were given in two blocks of ten, one block in the morning, one in the afternoon or evening. The two daily blocks of trials were separated by an interval of 4 hr. Within each block trials were separated by a 5-min. interval. Positive and negative trials within each block were given in a random order, varied from block to block except that: (1) Each block contained five positive and five negative trials. (2) Not more than three positive or three negative trials were given successively. (3) Over ten blocks of trials a positive was followed by a positive and a negative by a negative on 50% of trials. The positive shape remained in the water for 20 sec. on each trial (unless attacked sooner); if an octopus failed to attack within 20 sec., no reward was given and the shape was withdrawn. Since twice the number of trials per day were being given as in experiment I, the size of reward was reduced to a piece of sardine weighing 2.5 g. The negative shape always remained in the water for 20 sec.

The changes in method in this experiment were made for the following reasons: (1) Reward was made dependent on attack within the time limit in order to prevent octopuses learning not to attack until the sardine was presented with the shape. (2) Twenty trials a day were given in two groups of ten since this is a more convenient method of presenting trials from the experimenter's point of view, and Young (1959) has shown this method to be feasible for simple discriminations. (3) The period for which shapes remained in the water was reduced from 30 sec. to 20 sec., since few attacks are in any case made after 20 sec.

Results

The curves of learning for groups C, D and E are plotted in terms of percentage correct responses in Fig. 1. Group C (trained to discriminate upright and inverted $\sqcap$ figures) discriminated significantly better than either of the other two groups. The value of t for the difference between the scores of groups C and D over 12 days' training was 3.4 (D.F. = 14, $P < 0.01$), and for the difference between groups C and E (over 10 days' training only) 2.7 (D.F. = 14, $P < 0.02$). The difference between groups D and E was not significant ($t = 0.8$, D.F. = 14, $P > 0.4$). The performance of group D did not differ from chance, but group E discriminated better than chance over the last 6 days of training (χ^2 for the difference between the ratios attacks on positive to total attacks and attacks on negative to total attacks is 13.1, D.F. = 1, $P < 0.001$). For none of the three groups were there any significant differences between animals trained positive on one shape and animals trained positive on the other.

Thus the results of experiment II suggest that up-down mirror images are more readily discriminable by octopus than right-left mirror images, and that up-down mirror images are not particularly difficult to discriminate since the pair $\sqcap$ and $\sqcap$ was more readily discriminable than the pair $\sqcap$ and $\sqcap$.

In order to discover whether discrimination of the right-left mirror image $\sqcap$ shapes was possible, three animals from group D were given a further 18 days' training; three animals from group C were also given an additional 18 days' training in order to discover whether performance would improve further after the initial 240 trials. The performance of these six animals over the whole 30 days is shown in Fig. 4. The three animals from group D clearly learned the discrimination though to a low level of accuracy: over the last 300 trials they only performed with an accuracy of 54.5%. Both groups showed some improvement after the initial 12 days' training, and it is clear that 12 days is not sufficient for performance curves to reach an asymptote with a difficult discrimination and the method of training here used. Neither group showed any improvement in performance after day 20 (i.e. after 400 training trials).

The three animals from group C were given twenty transfer test trials at the conclusion

of their 30 days' training. In the transfer tests large figures were presented. The proportions of the figures used in the transfer tests remained the same as those of the original figures, but their areas were 2.25 times those of the original figures. No rewards or punishments were given in connexion with the transfer shapes. Accuracy of performance was 68%; since the percentage of correct responses given over the last 5 days of training had been 64%, this represents complete transfer.

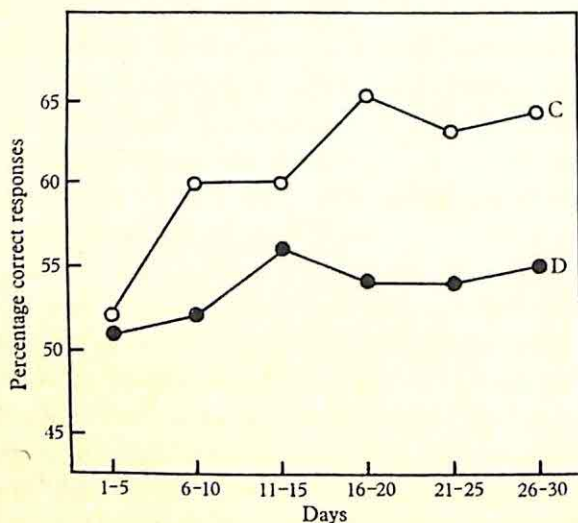


Fig. 4. Performance of three animals from group C, and three from group D over 30 days (600 trials). N.B. 50% represents the chance level of response.

Fig. 2 shows the way in which the total number of attacks made fell off for group C, when total attacks made are plotted as a percentage of total trials. It will be seen that the percentage of attacks given is actually less than that made by group A at least during the early part of training. This is presumably due to (1) The omission of reward in cases where an animal failed to attack within 20 sec. (Young (1958) has shown that feeding itself increases the tendency to attack). (2) The extra number of trials given a day, and the smaller time intervals between trials may also have contributed. Shock decreases the tendency to attack for some time after it is given, and this would presumably result in group C attacking less often than group A, since trials were more closely spaced for group C. Comparison of the curves for group C in Figs. 2 and 3 reveals that as discrimination improved, the total number of attacks given showed a tendency to increase. In fact in the group of three animals who were given a further 18 days' training no further increase in the daily number of attacks occurred: they continued to attack on only 25% of the total trials. It must be remembered that the rise in number of attacks exhibited by group A only occurred as a result of a change in the method of training and was not a consequence of learning to discriminate and thus avoiding shock on a higher proportion of attacks made. Fig. 3 shows that by the 12th day of training 76% of all attacks were being made by animals of group C on the positive figure. Whether performance is plotted as percentage of correct responses to total trials (Fig. 1), or as percentage of attacks on the positive shape to total attacks (Fig. 3) a continuous improvement in animals' discriminatory performance is shown, when the present training method is used.

Groups D and E showed a more marked falling off in level of attacks than group C. They were in fact only attacking on 10% of trials at the conclusion of discrimination training.

V. DISCUSSION

Theoretical implications

From the results of both experiments it is clear that octopuses can discriminate both right-left and up-down mirror images. In experiment II, however, animals trained on right-left mirror images only reached a significant level of discrimination when more than 12 days' training was given. In both experiments, animals trained on the up-down mirror images discriminated better than those trained on the left-right mirror images, though the difference was only significant in experiment II. The difference due to left-right and up-down mirror images being used was probably masked in experiment I because animals may have been learning not to attack the shape until the sardine was shown with it. It seems safe to conclude that for the octopus up-down mirror images are easier to discriminate than left-right mirror images.

In view of the octopus's ability to discriminate both up-down and left-right mirror images, hypothesis (1) as set out in the introduction must be rejected. It would appear that octopuses are able to analyse the relative extents of shapes along both the horizontal and vertical co-ordinates. However, they are able to discriminate better when the shapes they are discriminating differ in their horizontal extents than when they differ in their vertical extents. This means that any analysing system which operates in the octopus's brain must analyse horizontal extents of shapes more accurately than vertical extents of shapes. A neurological mechanism for analysing horizontal and vertical extents of shapes has in fact been suggested by the writer (Sutherland, 1957*b*). On the basis of the present experiments it would appear that the parts of the system responsible for analysing horizontal extents function more accurately than those parts responsible for analysing vertical extents.

On the basis of the analysing mechanism suggested various other predictions are possible, several of which have been tested (Sutherland, 1958*a*, 1959*a*, 1959*b*). One such prediction is of particular relevance to the present results. If the theory were correct, then not all mirror images should be discriminable. Any mirror images whose horizontal and vertical extents were the same should be indiscriminable. It has in fact been found that when such a pair of mirror images (a V and an inverted V) are used, octopuses are unable to learn to discriminate between them (Sutherland, 1959*b*). This result is particularly significant since the animals were trained by exactly the same method as that used in the present experiment II. Although the mirror images were up-down reversals, animals failed to achieve better than chance scores after 12 days' training in marked contrast to the animals trained in the present experiment to discriminate between up-down mirror image $\sqcup$'s.

The reason why up-down mirror images are more readily discriminated by the octopus than right-left mirror images may be that the equivalence of right-left mirror images has been learned since a right-left inversion of a shape can occur through the octopus swimming round to the other side of it, whereas this is not the case with up-down mirror images. In addition the vertical axis is more clearly differentiated by stable cues emanating from both within and without the organism than is the right-left axis. Thus more effort

has to be expended to move vertically than to move horizontally, and the direction 'up' is differentiated from 'down' by the water surface at the top and floor of the tank (or of the sea) at the bottom, and by differences of illumination. The fact that learning may play a part in producing greater ease of discrimination for shapes differing in their horizontal projections than for shapes differing in their vertical projections, does not mean that the system operating is not that sketched above: it would merely mean that learning played some part in the setting up of the system.

The finding that a $\sqcup$ on its side is more difficult to discriminate from an upright or an inverted $\sqcup$, than are the latter figures from one another is neither for nor against the theory as so far expounded. It does suggest, however, that mirror image shapes are not per se particularly difficult for the octopus to discriminate, and it lends no support to theories which predict that all mirror images should be poorly discriminated (e.g. Deutsch, 1955). Any adequate theory of shape discrimination must explain why some mirror image figures are difficult or impossible to discriminate (Sutherland, 1957*a*, 1959), while others are readily discriminable.

Comparison of results with other data

There is no evidence to suggest that animals have any special difficulty in discriminating mirror images as such—except mirror images of oblique lines (Sutherland, 1957*a*). Lashley (1930, 1938) has shown that rats can discriminate upright and inverted triangles and also triangles with apex facing left and apex facing right. Similarly he found, rats could discriminate upright and inverted $\sqcup$ figures of the type used in the present experiment and $\sqcup$ figures lying on opposite sides. Rats discriminated more readily between the up-down reversed $\sqcup$'s than between the right-left reversed ones, but because of the rat's tendency to discriminate only the part of the figure near the base line it is difficult to interpret these findings. Kirk (1936) showed that rats could discriminate 'F' from its right-left mirror image. Lashley (1938) comes to the conclusion that 'The learning scores with symmetrical figures do not significantly exceed those for non-symmetrical figures of the same order of complexity.'

Karn & Munn (1932) found that dogs could discriminate an upright from an inverted triangle, and since they transferred the discrimination to patterns of three dots they must have been discriminating in terms of the whole figure and not merely in terms of the base line.

Harlow (1945) found that rhesus monkeys discriminated less readily between eight pairs of mirror image figures than between certain other pairs of figures (e.g. a triangle and a square). When, however, one member of each pair of mirror images was arbitrarily combined with one member of another pair of mirror images, he found that the difficulty of discrimination was still as great as for the original mirror image pairs. He also found that monkeys discriminated more readily between a $\top$ and an inverted $\top$ than between two $\top$'s laid on opposite sides.

Thus the available evidence for mirror image discrimination in mammals is in full agreement with the results of the present experiments on the octopus: mirror images per se are not difficult to discriminate, and up-down mirror images tend to be more readily discriminable than right-left mirror images.

Methods

It has already been necessary to discuss at some length the methods of training in presenting the results. For convenience, however, the main conclusions about methods will be summarized here. (1) In training octopuses on a difficult discrimination, the method of presenting a sardine in association with the shape on *every* positive trial produces rapid initial learning, but performance fails to improve after the first twenty trials. It seems likely that animals are learning not to attack the positive figure until the sardine is presented with it, and as a consequence discriminatory performance fails to show any improvement. (2) If reward is made contingent on attack the initial rate of learning is slower, but improvement is more continuous and it seems likely that a higher final level of performance is reached. That initial learning is slower seems certain on the basis of the present experiment, since the group trained on up-down inverted $\sqcup$'s learned slower than that trained on up-down inverted $\sqcap$'s, although the final level of performance of the former group was higher than that of the latter. This is corroborated by a further experiment (Sutherland, 1958c) in which it was found that animals trained to discriminate between a circle and a square learned more slowly if the reward was made contingent on attack on the positive figure alone than if reward was presented with the shape on positive trials on which the animal failed to attack within 30 sec. The conclusion that a higher level of correct responses is reached where reward is made contingent upon attack is open to doubt, since it is not certain whether the upright and inverted $\sqcap$ figures are as readily discriminable as the upright and inverted $\sqcup$ figures. (3) With both methods of training there is a marked fall in the total number of attacks made as training proceeds. Initially the percentage of attacks made is lower with the method of making reward contingent on attack on the positive shape presented on its own, but with this method, some recovery occurs as discrimination improves (cf. Sutherland, 1959b). Since the percentage of attacks made correlates with the ease of discriminability of the shapes, it does not matter whether we use a score based on the ratio correct responses to trials or a score based on the ratio attacks on positive to total attacks, if we merely wish to establish the relative discriminability of different pairs of shapes. In fact the percentage of correct responses, although giving a figure which underestimates the accuracy of the discrimination, is the more convenient score for comparative purposes, since it is easy to ensure that different animals are given the same number of trials but difficult to ensure that they complete the same number of attacks. (4) On the whole the method of giving twenty trials a day and making reward contingent on attack would appear to secure the best results, since (a) animals cannot learn to gain reward by waiting until the sardine is presented with the shape before attacking; (b) a smoother learning curve is obtained by this method and final performance is probably higher; (c) although more trials a day are given, this method of giving trials is considerably easier from the experimenter's point of view. Giving trials at closely spaced intervals avoids the labour of securing the lids of the octopus's tanks between each trial in order to prevent them escaping. The chief disadvantage of this method is that many trials are wasted since animals cannot learn anything on any trial on which they fail to attack within the time limit, whereas by rewarding every trial they might learn to associate the positive shape with food even on positive trials where they fail to attack within the time limit.

The writer wishes to thank the Royal Society for a grant which made the experimental work possible, the Director and staff of the Stazione Zoologica, Naples, for the hospitality and facilities afforded him there, Professor J. Z. Young, F.R.S., for advice and encouragement, and Mr W. R. A. Muntz for help in conducting the experiments. The writing up of the results was facilitated by a grant from the Nuffield Foundation for a project on stimulus analysing mechanisms.

REFERENCES

- BOYCOTT, B. B. (1954). Learning in *Octopus vulgaris* and other Cephalopods. *Pubbl. Staz. Zool. Napoli*, **25**, 67-93.
- DEUTSCH, J. A. (1955). A theory of shape recognition. *Brit. J. Psychol.* **46**, 30-7.
- HARLOW, H. F. (1945). Studies in discrimination learning by monkeys: III. Factors influencing the solution of discrimination problems by rhesus monkeys. *J. Gen. Psychol.* **32**, 213-27.
- KARN, W. H. & MUNN, N. L. (1932). Visual pattern discrimination in the dog. *J. Genet. Psychol.* **40**, 363-74.
- KIRK, S. A. (1936). Extra-striate functions in the discrimination of complex visual patterns. *J. Comp. Psychol.* **21**, 145-59.
- LASHLEY, K. S. (1930). The mechanism of vision: I. A method for rapid analysis of pattern vision in the rat. *J. Genet. Psychol.* **37**, 453-60.
- LASHLEY, K. S. (1938). The mechanism of vision: XV. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **18**, 123-93.
- SUTHERLAND, N. S. (1957*a*). Visual discrimination of orientation by octopus. *Brit. J. Psychol.* **48**, 55-71.
- SUTHERLAND, N. S. (1957*b*). Visual discrimination of orientation and shape by octopus. *Nature Lond.*, **179**, 11-13.
- SUTHERLAND, N. S. (1958*a*). Visual discrimination of shape by octopus: Squares and triangles. *Quart. J. Exp. Psychol.* **10**, 40-7.
- SUTHERLAND, N. S. (1958*b*). Visual discrimination of the orientation of rectangles by *Octopus vulgaris* Lamark. *J. Comp. physiol. Psychol.* **51**, 452-8.
- SUTHERLAND, N. S. (1959*a*). Visual discrimination of shape by octopus. Circles and squares, and circles and triangles. *Quart. J. Exp. Psychol.*, **11**, 24-32.
- SUTHERLAND, N. S. (1959*b*). A test of a theory of shape discrimination in octopus. *J. Comp. physiol. Psychol.*, **52**, 135-41.
- YOUNG, J. Z. (1958). Responses of untrained octopuses to various figures and the effect of removal of the vertical lobe. *Proc. Roy. Soc. B*, **149**, 463-83.
- YOUNG, J. Z. (1959). Effect of various rates of training and its reversal on octopus with and without vertical lobes. *Proc. Roy. Soc. B*, (in the Press).

(Manuscript received 25 September 1959)

A GROUP VERSION OF A PERCEPTUAL MAZE TEST

BY ALICK ELITHORN,* MYFANWY KERR* AND JOY MOTT†

From the Medical Research Council Neurological Research Unit and the Institute of Neurology, National Hospital for Nervous Diseases, Queen Square, London, and the Royal Free Hospital, Gray's Inn Road, London

Two group versions of a multiple-choice perceptual maze test are described. The test has been administered to two groups of normal subjects. The results show that the test has a high reliability and suggest that the material reliably samples an ability which is not normally measured by intelligence tests. It is tentatively suggested that this ability may be related to perceptual capacity. The importance of perceptual factors in determining intelligent behaviour is briefly discussed.

I. INTRODUCTION

Functions which are impaired by frontal lobe damage may be tested by pencil and paper maze tasks. This was pointed out by Porteus & Kepner (1944) and has been repeatedly confirmed, most recently by A. Smith (personal communication, 1957).

A previous paper (Elithorn, 1955) described a maze task for which the material was so designed that some of the factors which determine performance could be systematically varied. A preliminary investigation indicated that test performance was sensitive to frontal and temporal lobe injury. Test scores were found to correlate positively with post-operative social adjustment, verbal and performance intelligence test scores and negatively with chronological age.

In the present study results obtained with normal subjects with a group version of this test are reported.

II. MATERIAL AND METHODS

The test material consists of patterns each of which has a lattice background of paths or tracks represented by dotted lines. At a variable number of intersections larger dots (solid black circles) are superimposed. The subject is required to find in each pattern a pathway from bottom to top which passes through the greatest possible number of the circles. There are two main restrictions. First, the subject must keep to the paths or tracks and must not cut across from one path to another. Secondly, at any junction point the pathway chosen by the subject must continue forwards, i.e. it may fork left or right, but must not double back. Each pattern is so designed that there is only one correct series of circles through which the pathway can pass. In order that the subject may know when he has found a correct pathway the maximum number of circles which can be obtained is printed below each pattern.

In the present investigations the paths were represented by finely dotted lines and the large dots at the intersections were solid black circles (Fig. 1). These patterns were printed on the second, third and fourth pages of two booklets. They were arranged in an approximate order of difficulty. Booklet A contained the first thirty patterns, booklet B the remaining twenty. In order to provide an alternative version of the test two additional booklets were prepared—A(1) and B(1)—in which the patterns were inverted. This provided patterns of comparable difficulty, but prevented subjects from benefiting from possible recall of previous solutions. In Fig. 1 pattern (b) is pattern (a) inverted. The first page of each booklet was identical and consisted of short instructions and five examples with solutions.

The following were the instructions:

'This test consists of a number of maze puzzles. In each one you have to find and mark with a pencil

* Supported by the Medical Research Council.

† Supported in part by the Medical Research Council and in part by the Royal Free Hospital Endowment Fund.

the path from bottom to top which passes through the most dots. Below each puzzle is a number like this [3] which tells you how many dots you can get. When you find a path with this number of dots, mark it and go on to the next puzzle. Do not waste too much time on any one puzzle—you may come back later. Read the following rules carefully.

- (1) You may start anywhere along the bottom line of each puzzle.
- (2) You must keep to the lines and not jump across the spaces.
- (3) You must always go forward. At each junction you may take the left or right path, but you must not double back.

The two identical puzzles to the right show you how this is done. Look at them and then try the ones below, marking in the paths with your pencil.

(a) Investigation I

(i) Subjects

One hundred and ninety National Servicemen acted as subjects and ninety-nine of these subjects were retested fourteen days later. The sample made available consisted of recruits who had initially been selected as 'potential officer material'. From this original group subsequent selection procedures had removed most of those finally chosen for officer training. This selection meant that any correlations would be attenuated by the homogeneity of the group for general ability. It was, however, decided to test this non-representative sample because it was the only group available at that time for whom it was possible to obtain scores on several other tests and because it was the only group in which a reasonable number of subjects would be available for retest.

(b) Investigation II

Since very encouraging results were obtained in the first investigation, a further sample more representative of the general population was subsequently made available for testing. This second group consisted of 226 unselected recruits. It is probable that defective and illiterate individuals had been excluded by preliminary National Service Boards; further, a few more intelligent subjects avoid recruit intake selection procedures by making a direct entry to commissioned rank from an O.T.C. or as doctors. With these exceptions the second group was a representative and unselected sample of the population liable for National Service.

In both investigations the scores of each subject on a battery of other selection tests were made available. This battery included tests of mechanical and electrical knowledge; a verbal intelligence test; a non-verbal intelligence test; tests of spatial ability; spelling; and arithmetic and algebra. These tests will be referred to as RC1 (m) and RC1 (el); RC2; RC3; K1; V1; and N, respectively.

(a) Investigation I

(ii) Test and retest procedures

One hundred and ninety subjects were tested. They were subdivided into two groups depending on the test material used. Group I consisted of 142 subjects and group II of 48 subjects. The test-retest procedure was as follows:

Test form used	Test		Retest		Not retested
	A and B	A1 and B1	A and B	A1 and B1	
Group I	142	—	25	30	87
Group II	—	48	21	23	4

At the retest session the subjects were not told whether they were doing the same or a different version of the test. They were simply asked to do some more maze patterns.

Since it was not known what the optimal test duration would be, the time allowed for each booklet was divided into two. During the first 10 min. each subject used a red pencil, and for the next 5 min. a blue one. After 15 min. with booklet A (or A1) the subjects were told to stop and the papers collected. Booklet B (or B1) was then distributed and the same procedure followed with two periods of 10 and 5 min.

During the retest only the initial 10-min. period was allowed for the A booklet. Ten minutes, with a subsequent 5 min., was allowed as before for booklet B.

(b) Investigation II

Because of shortage of time only booklet A could be administered and the testing time was 10 min.

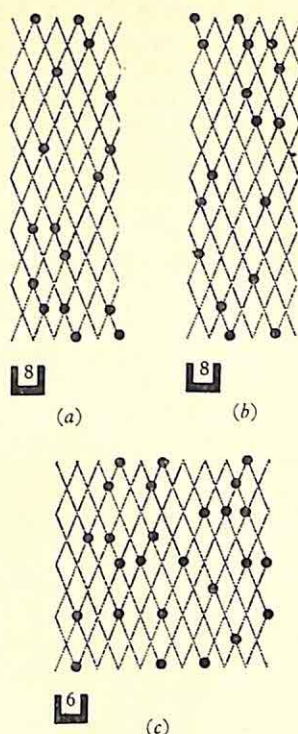


Fig. 1. (a) Test item from booklet A; (b) test item from booklet A1 (the alternative version), it is (a) inverted; (c) test item from booklet A indicating how the size of the units may be varied.

(iii) Results

(a) Investigation I

Figs. 2a and b show the scores obtained with the 142 subjects who did the A and B version of the test. Figure 2a shows the distribution of scores (hatched shading) obtained when only the first 10 min. of each period were scored. Solid shading indicates the scores obtained with the full 15-min. testing period. In Fig. 2b the results for booklet B are similarly indicated.

Because of the selected nature of the subjects in this investigation, care must be taken in interpreting the distribution of scores. It is perhaps worth noting that out of 190 subjects only one completed booklet A (or A1) in 10 min. (mean score for A group 16.11 ± 0.4). In 15 min. fourteen subjects completed book A (or A1), and this length of testing therefore led to a loss of discrimination with this group (mean score for A group 20.75 ± 0.48).

The distributions of scores for booklet A are skewed to the right. Since a similar distribution was found for the scores with RC2 and RC3 which are known to have a normal distribution with unselected populations it seemed likely that much of this effect might be related to the fact that the group had been initially selected for intellectual ability. The observed means for RC2 and RC3 were 19.4 ± 0.48 and 19.0 ± 0.42 ($n = 142$). These means are markedly different from the established means for the population of their age-group (15.0, s.d. 6.67 and 15.3, s.d. 5.99).^{*} These differences are highly significant and confirm that the present sample is well above average in intelligence.

With booklets B and B1 approximately half the twenty items were, on the average, successfully solved in 10 min. (mean score for B group 9.61 ± 0.27), only two additional items being solved in the extra 5 min. (mean 15-min. score for B group 11.65 ± 0.29). It seems likely, therefore, that for a timed test the increase in difficulty of the individual items is too steep in these books. The results with the equivalent forms and on retesting are discussed below.

* These figures were kindly supplied by Mrs Anthony.



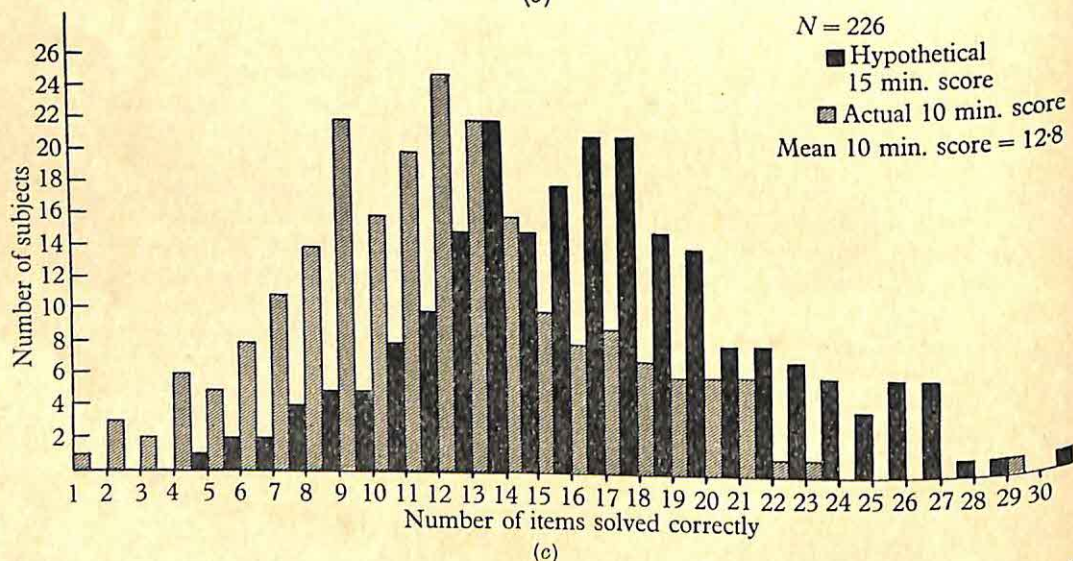
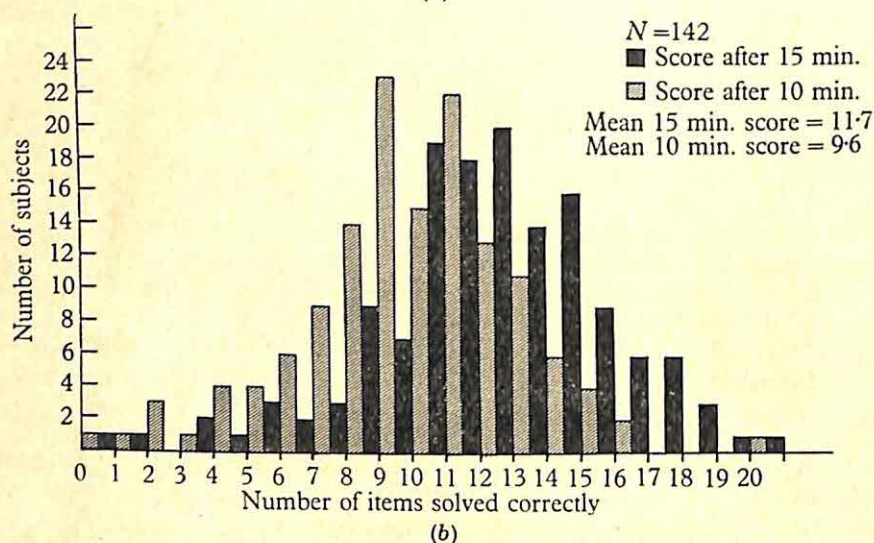
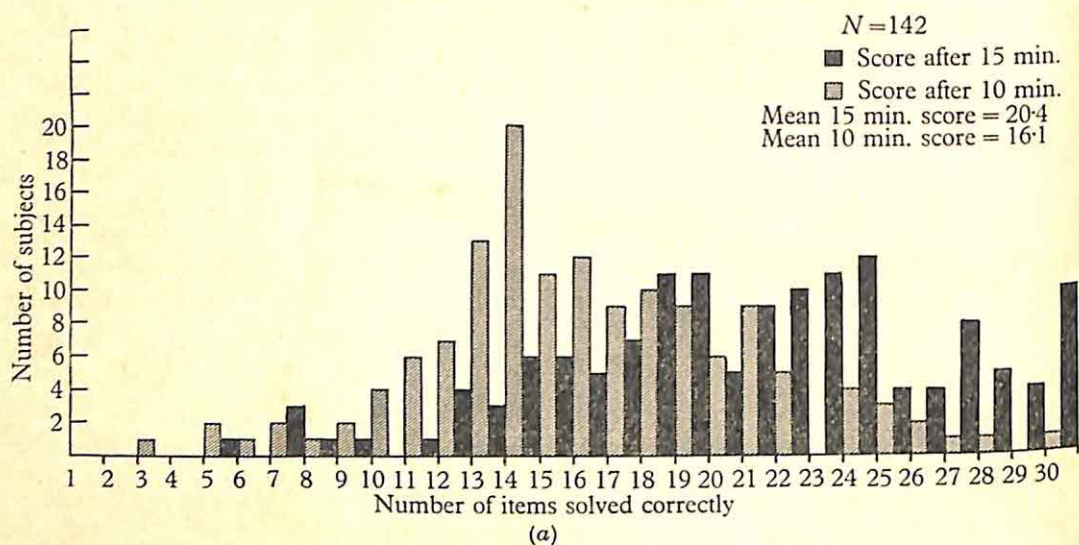


Fig. 2. (a) Investigation I. Maze score distribution. Test booklet A only. (b) Investigation I. Maze score distribution. Test booklet B only. (c) Investigation II. Maze score distribution. Test booklet A only.

(b) Investigation II

Fig. 2c gives the distribution of 10-min. scores for booklet A (hatched). As explained above, it was anticipated that this group would be a normal sample of the population liable for military service. The scores obtained with RC2 and RC3 confirm that this was the case. The observed mean obtained with RC3 does not differ significantly from the established norm 15.3 (s.d. 5.99). The observed mean obtained with RC2, 14.0 ± 0.35 , is just significantly different ($p < 0.05 > 0.01$) from the established norm of 15.0 (s.d. 6.67). The median number of mazes correctly solved was 12; the mean was 12.8 with a standard deviation of 4.97. The other coefficients of the distribution were: $g_1 = +0.27$, $t = 1.67$, $p < 0.1 > 0.05$, $g_2 = +0.32$, $t = 0.98$, $p < 0.4 > 0.3$. Skewness and kurtosis, therefore, do not depart significantly from normal. Only 31% of the subjects solved half the items and 9% solved two-thirds. This marked clustering of scores at the lower end of the scale implies a considerable loss of discrimination, and suggests that although 15 min. was too long for the selected sample tested in the first investigation 10 min. is inadequate for an unselected sample. As mentioned, it was impracticable to obtain 15-min. scores from this group, but it seems likely that the longer time would not lead to the loss of discrimination found in the first investigation.

It is of practical interest to decide whether it is likely to be more profitable to repeat the present test with a 15-min. testing time or whether the item contribution of the test should first be altered. The results obtained in the first investigation were, therefore, used to extrapolate from the present scores. This transformation, which is theoretically indefensible, gives a hypothetical 15-min. distribution represented by the solid columns in Fig. 2c. Further analysis is not justified but this approximation would suggest a further trial of booklet A with a 15-min. testing period.

(iv) Test-retest reliability, practice effects, alternative version

In the first investigation ninety-nine subjects were retested 14 days after the initial test. Of these forty-eight were given the same version of the test as they had previously used. Fifty-one were given the alternative version. The correlation coefficient for the test-retest scores (A and B 10 min. only) with

Table 1. Investigation I

Test				Retest			
Form	N	10 min. score	15 min. score	Form	N	10 min. score	15 min. score
Group I							
A	55	15.91 ± 0.58	20.31 ± 0.66	A	25	23.8 ± 0.92	—
				A1	30	22.4 ± 0.84	—
B	55	9.71 ± 0.39	11.56 ± 0.46	B	25	12.68 ± 0.62	14.52 ± 0.61
				B1	30	11.93 ± 0.56	14.03 ± 0.61
Group II							
A1	44	16.36 ± 0.79	21.02 ± 0.99	A	21	23.62 ± 1.03	—
				A1	23	22.09 ± 1.15	—
B1	44	10.09 ± 0.52	12.16 ± 0.57	B	21	12.52 ± 0.74	15.24 ± 0.69
				B1	23	11.78 ± 0.96	13.69 ± 0.91

Mean scores and standard errors of Groups I and II on the maze test and retest.

identical material was 0.89 ($n = 48$) and for the test-retest scores when retesting with an alternative version $r = 0.81$ ($n = 51$). These results compare very favourably with the test-retest results reported for the individual items of the WAIS battery (Wechsler, 1955). Only five subtests have reliabilities greater than 0.85; three of these are in the verbal scale—information 0.91, similarities 0.87 and vocabulary 0.94—and two in the performance scale—digit symbol 0.92 and block design 0.86. Test-retest correlations of the Progressive Matrices range between 0.76 and 0.91 (Foulds, 1947; Vernon & Parry, 1949).

As would be expected, the test-retest correlation was higher when the retest was undertaken with identical material than it was when fresh material of a similar character was presented. This difference ($r = 0.89$ and 0.81 respectively) was not statistically significant and might perhaps have been a chance finding ($p < 0.2 > 0.1$). Further evidence that some recall affected the test-retest correlations was provided by the observation that the increase in the number of items solved was greater when these

were the same at the second presentation (mean gain: 9.9 ± 0.83 , $n = 48$) than when different items were presented (mean gain: 8.5 ± 0.85 , $n = 51$). This difference in the effect of practice in the two situations is in the expected direction, but is not marked and is not statistically significant (see Table 1).

It can be tentatively concluded, therefore, not only that the reliability of the test is high, but also that where the same items have to be re-used the recall of individual solutions does not contribute much to the practice effect or to retest correlations.

Table 2 is a correlation matrix for the tests used. A preliminary analysis suggests that the correlations obtained between the maze and the other tests in the battery could largely be explained by two general factors. Of these one would be a factor of general intellectual ability. The second would be a bipolar educational/non-educational factor. This analysis supports the view put forward below that the maze material tests an aspect of non-cultural intelligence that is not adequately sampled by conventional intelligence tests. Thus although its reliability is high the correlations between the maze scores and those of the other tests designed to test intelligence are relatively low. One of the tests, RC2, is a verbal test of general intelligence and the other is a non-verbal test similar in character to the progressive matrices. When the scores of both these tests are combined (investigation I) the correlation with the maze scores is not appreciably higher ($r = 0.46$).

Table 2. *Correlation matrix for test battery*

	RC1 (m)	RC1 (el)	RC2	RC3	K1	V1	N
RC1 (m)	—	—	—	—	—	—	—
RC1 (el)	0.64	—	—	—	—	—	—
	<i>0.66</i>	—	—	—	—	—	—
RC2	0.23	0.41	—	—	—	—	—
	<i>0.26</i>	<i>0.36</i>	—	—	—	—	—
RC3	0.32	0.47	0.69	—	—	—	—
	<i>0.30</i>	<i>0.31</i>	<i>0.64</i>	—	—	—	—
K1	0.40	0.35	0.44	0.47	—	—	—
	<i>0.26</i>	<i>0.19</i>	<i>0.19</i>	<i>0.35</i>	—	—	—
V1	0.18	0.35	0.69	0.54	0.29	—	—
	<i>0.20</i>	<i>0.30</i>	<i>0.67</i>	<i>0.47</i>	*	—	—
N	0.22	0.45	0.85	0.69	0.33	0.65	—
	<i>0.26</i>	<i>0.30</i>	<i>0.77</i>	<i>0.57</i>	*	<i>0.57</i>	—
Maze	0.27	0.30	0.41	0.44	0.41	0.28	0.37
	<i>0.08</i>	<i>0.18</i>	<i>0.46</i>	<i>0.37</i>	<i>0.31</i>	<i>0.45</i>	<i>0.44</i>

* Data not available.

Correlations from investigation I (booklet A and B 15-min. scores $n = 142$) in roman type; from investigation II (booklet A 10-min. scores $n = 226$) in italics.

RC1 (m) and RC1 (el) were tests of mechanical and electrical knowledge respectively; RC2 a verbal intelligence scale; RC3 a non-verbal abstraction test; K1 a spatial orientation test; V1 a spelling test; N a test of arithmetical and algebraic abilities.

III. DISCUSSION

The test-retest reliabilities (0.89 for the same version of the test, 0.81 for the alternative version) obtained with normal subjects show convincingly that this test material is sampling in a reliable way a skill which is consistently different for different individuals.

The results also confirm that the scores correlate positively and significantly with measures of general intelligence. In the present data there are not sufficient correlations with other tests to justify a factorial analysis, but some preliminary calculations suggested that two general factors of intelligence would account for most of the inter-correlations of the maze test and other tests in the battery. These factors—a general intelligence factor and a bipolar academic/non-academic factor—indicate that the maze is sampling chiefly non-cultural aspects of intellectual functions.

In the previous study, maze test performance correlated highly ($r = 0.74$) with the Kohs' blocks item of the Wechsler Adult Intelligence Scale. In the present investigations the correlations with the test designed to test spatial abilities (test K) was low ($r = 0.41$).

and 0.39 respectively), in fact, lower than the correlations with the verbal test of intelligence RC2 ($r = 0.41$ and 0.46 respectively). Part of the discrepancy may be due to the fact that the original correlation obtained with the Kohs' test was, as then suggested, exaggerated by the inclusion of subjects with focal brain damage, and should therefore be interpreted with caution. An additional explanation may be that the K* test used in the present investigation was later found to be unreliable. Further, Vernon and Parry have noted that K tests probably test several different though related aspects of spatial ability.

In the previous report it was suggested that once the instructions and the method of solving the maze problems had been appreciated the solution of additional similar items required no further logical reasoning and that success or failure was determined mainly by a perceptual skill. It was further suggested that complex visual tasks of this nature might be dealt with by a perceptual scanning mechanism and that differences in performance might reflect differences in the size of the perceptual unit which any particular individual can effectively handle and the rate at which such units can be handled.

While speed and power are two aspects of intellectual function which are accepted as being closely correlated, a great variety of opinion exists among psychologists as to whether they have an independent existence and as to their relative importance in determining intellectual ability. However, present-day thought inclines to the view that such a dichotomy applies only to intelligence test administration and not to the functions being measured. Spearman & Wynn Jones (1950) while unable to reject the dual concepts completely remark that 'experimental evidence contradicts the existence of any group factor extending over the whole sphere of cognizing quickly'.

With the present test it is relevant to consider that the maximum size of the perceptual unit with which the individual can deal effectively might very well not be directly dependent on the maximum rate at which such units can be handled. It has been suggested that optimal performance is obtained by a perceptual scanning rather than by a rigid sequential search. Scanning in this context implies that each perceptual unit is considered in isolation and then integrated. There are at present no experimental data on this subject, but subjective reports indicate that perceptual scanning is the method of choice but that this may give way to a sequential search when it has failed to reveal the correct path.

The suggestion by Cattell (1943) that 'non-cultural' or 'fluid' intellectual ability is partly dependent upon, or at least closely correlated with, the capacity to discriminate relationships is of interest here. It could perhaps be related to Lashley's distinction between perceptual difficulty and logical difficulty. Lashley (1938) suggested that 'identical reaction to the identical existential elements in different situations is logically the simplest form of generalized reaction. That it is physiologically the simplest or most primitive is doubtful'.

A basic criticism of intelligence tests is that they are pragmatic and relate intelligent behaviour to the ability to reason logically. The syllogism, like the wheel, not only produces a qualitative change in behaviour, but is a cultural achievement; neither is likely to reflect directly a fundamental property of nervous tissue. If the suggestion that the present material tests 'perceptual capacity' can be confirmed, then the correlations observed with intelligence tests (RC2 and RC3) in the battery used would support

* The test is similar in construction to NIIP test 94.

Lashley's view that perceptual skills are phylogenetically fundamental components of intelligent behaviour.

Porteus claims that scores on his maze test are, in some ways, a more sensitive index of social adaptive capacity than other tests of 'intelligence'. Teuber (1950) has found that a 'field of search' test discriminates patients with frontal lobe damage from controls, and emphasizes the importance of perceptual factors in determining 'intelligent' behaviour. Similar observations have been made by Halstead (1947). There is much to be said for concepts of intelligence which relate intelligent behaviour to our knowledge of cerebral organization.

Further investigation of the present material would be justified if it were possible to control systematically the perceptual complexity of individual patterns. Movable type which permits this has been developed and will be described in a subsequent paper.

It would not have been possible to undertake these group tests without the encouragement and help of Dr John Parry. We are also greatly indebted to Miss Mary Allan and Mrs S. Anthony for help in the administration and analysis. Mrs C. M. Lawrence helped with the administration and scoring of the tests. Mrs M. A. Beale and Mr A. H. Prickett prepared the diagrams from which the plates for the test booklets were prepared.

The views expressed do not in any way represent those of the Service department to which we are indebted for access to our subjects.

REFERENCES

- CATTELL, R. B. (1943). The measurement of adult intelligence. *Psychol. Bull.* **40**, 153.
- ELITHORN, A. (1955). A perceptual maze test sensitive to brain damage: a preliminary report. *J. Neurol. Neurosurg. Psychiat.* **18**, 287.
- FOULDS, G. A. (1947). Report of an experimental survey. In *Progressive Matrices*. Dumfries: Crichton Royal.
- HALSTEAD, W. C. (1947). *Brain and Intelligence*. Chicago: University of Chicago Press.
- LASHLEY, K. S. (1938). The mechanism of vision. 12. Preliminary studies of the rat's capacity for detail vision. *J. Gen. Psychol.* **18**, 123.
- PORTEUS, S. D. & KEPNER, R. de M. (1944). Mental changes after bilateral prefrontal lobotomy. *Genet. Psychol. Monogr.* **29**, 4.
- SPEARMAN, C. & WYNN JONES, L. (1950). *Human ability*. London: Macmillan.
- TEUBER, H. L. (1950). Neuropsychology. In *Recent Advances in Diagnostic Psychological Testing*. Springfield: Thomas.
- VERNON, P. E. & PARRY, J. (1949). *Personnel Selection in the British Forces*. London: University of London Press.
- WECHSLER, D. (1955). *WAIS Manual*. New York: Psychological Corporation.

(Manuscript received 2 October 1958)

VISUAL SIZE AND FAMILIAR SIZE: INDIVIDUAL DIFFERENCES*

By A. C. McKENNEL

Psychology Department, Glasgow University

Recent experiments show that, in the absence of distance cues, accurate estimates of the size of familiar objects can be achieved on the basis of memory alone. It is by no means certain, however, that memory is operative in size estimates of familiar objects seen in ordinary perceptual situations, since distance cues normally present are themselves sufficient to mediate the perception of the real sizes of things. In investigating this problem use is made of the characteristic individual differences that occur when subjects are obliged to make use of visual cues to size (i.e. in size estimates of non-familiar objects). There is also a considerable variance in individual estimates from memory of the size of a familiar object. For a group of subjects data on these two sources of variance can be related to the variance in their visual estimates of the size of a common object. By this means the relative weights accorded to past experience with the object and to present visual cues can be determined. Results suggest that prior knowledge of familiar 'homogeneous' objects (occurring environmentally mainly in one standard size) entirely determined the estimate of their size, while even for less familiar and for 'heterogeneous' objects there is evidence, in some cases, that non-visual factors were partly active.

I. INTRODUCTION

The size constancy of familiar objects raises in a critical fashion an aspect of the more general and long-standing problem of the role of past versus present experience in perception. Recent experiments, referred to below, suggest that our impression of the constant, tangible size of familiar things may be a phenomenon quite distinct from the size constancy demonstrated for unknown objects introduced for the first time in laboratory experiments. Accurate estimates of the size of distant, non-familiar objects are necessarily mediated by the various binocular and/or monocular cues to distance, coupled with retinal image size. Size estimates of distant, *familiar* objects, however, are not necessarily based on the use of distance cues; they could also be based on the observer's prior knowledge of the size of the objects.

The issue is well brought out in a recent follow-up (Bolles & Bailey, 1956) of the original investigation by Brunswick (1944) of size estimates of everyday objects seen in natural surroundings. Brunswick found very high size constancy, which he interpreted as evidence of 'the distal focusing of perception'. It is this interpretation of the results which has since been disputed, since the experiment leaves unanswered the question how far familiarity with the everyday objects entered into the size estimates. Bolles & Bailey (1956) have carried out what was essentially a repetition of the Brunswick experiment, except that prior to each visual judgement they obtained the subject's non-visual estimate, from memory, of the size of the object. In 52% of the cases, subjects were not able to improve their estimates with the additional visual information; and for objects which could be adequately identified, judgements from memory were as accurate as visual judgements. We have, therefore, to consider in addition to the distal focusing of perception, the possibility of a 'central' focusing, in which retinal size and the retinal

* An extended version of a paper read at the Annual Conference of the British Psychological Society, Birmingham, 1958.

image are cues only to the identity of the object, the information on its real size being supplied by prior experience. The available evidence, however, provides no basis for deciding which of these explanations is the correct one. Since familiar objects normally occur in surroundings which provide adequate distance cues, either vision or memory or both could convey the information on real size. There is a need for investigation to determine what actually occurs.

A similar qualification applies to the demonstrations of the role of familiar or assumed size by the Transactionalist group (Kilpatrick, 1952). The Thereness-Thatness Table (Ittelson, 1951), among other demonstrations, shows quite dramatically that in certain situations the known or assumed size of an object can not only be utilized but will operate to determine the apparent location of the object in space. But the contrived stimulus arrangement on which these demonstrations depend serves to limit inferences from the results. In particular, since distance cues normally present are either eliminated or reduced, we are not bound to accept that assumptions from past experience operate in the same manner for objects seen in natural surroundings with adequate distance cues. As Allport (1955, p. 284) has stated about his own critique of these demonstrations, 'It is not implied that assumptions are not active in veridical perception, that is, in cases in which there is no hitch, but only that some more positive way of demonstrating them is needed'.

The crucial experiment here has to demonstrate not merely that individuals can use memory standards, but that they do so in situations where visual cues have an equal chance as potential determinants of the perception. This requirement raises basic difficulties for the customary approach to perceptual problems. In the absence of trustworthy introspective information, experimental psychologists have proceeded by obtaining the quantitative results of perceptual judgements, the inference to the basis of judgement being made after systematically varying or eliminating the various cues in order to see which are utilized. The present problem requires however that neither the memory nor the visual cues be eliminated from the judgement situation. Systematic variation of the amount of prior experience, while a valuable general approach to the effects of learning on perception, can hardly be applied to the size constancy problem, since the initially high degree of perceptual achievement here leaves little room for practice effects. It was in the light of these difficulties that the present, new approach utilizing individual differences in size constancy was adopted. Individual differences in perceptual constancy were first fully demonstrated by Thouless (1932), and have been studied by him and by later investigators as possible indicators of wider personality differences. In the present study, however, the interest is not so much in the individual differences for their own sake, as in the convenient sources of variation they provide in the main variables of the problem. A significant variation in the memory standards of familiar size employed by different individuals has been previously found by McKennell (1953, 1959). The existence of these sources of variation on both the memory and the visual sides suggested the present covariance approach to the study of the size constancy of familiar objects.

II. METHOD

The thirty-eight subjects in the main experiment were seated together at distances from a display screen varying between 7 and 20 ft. Three kinds of judgements, to be referred to as X_1 , X_2 and Y , were obtained by a drawing method as follows.

Judgements X_1 . Non-visual estimates of familiar objects. Here the subjects were instructed to draw a set of lines of which the lengths would represent their estimates from memory of the sizes of nine common objects, namely a matchstick, a 9-inch rule, a playing-card, a cigarette, a nail-file, a pocket-comb, a medicine-bottle, a cigar and a penknife. Each line was drawn on a separate sheet of paper, and the entire series of objects was judged twice, the second time in reverse order.

At the end of the experiment the subjects were asked to rate each of the common objects on a scale from zero to five for degree of *manipulative familiarity*, that is the relative amount of handling experience they had had with similar objects. In the case of the 9-inch rule it was agreed, after discussion, that ratings should correspond to the amount of general experience with inches as measurement units.

Four of the objects, the matchstick, the 9-inch rule, the playing-card and the cigarette, occur solely or mainly in one standard size, and, following the terminology introduced by Bolles & Bailey (1956), will be referred to as 'homogeneous' objects. The remaining objects each occur in a relatively wide range of sizes and will be described as 'heterogeneous' in size. For the heterogeneous objects, as for the others, the subjects were required to estimate 'the size most commonly met with'.

Judgements Y . Visual estimates of familiar objects. Subsequent to the memory judgements, X_1 , specimens of the actual objects were fixed in the centre of the display screen one at a time, while the subjects again drew a line to represent the length of each object. As before the lines were drawn separately on sheets of paper, one for each object, and the entire series of objects was presented for judgement twice.

Judgements X_2 . Visual estimates of non-familiar objects. These were judgements of the size of three squares, cut from white card, having sides 3.5, 5.5 and 8.5 in., to be referred to as squares j , k and l , respectively.

The memory judgements X_1 were made first. The initial visual judgements were then made on the squares j , k and l , placed in the display screen one at a time. These judgements, X_2 , were followed by judgements, Y , of the familiar objects. Squares j , k and l , were then presented for the second time in reverse order, with the second series of Y judgements following. For all visual judgements, Y and X_2 , the subjects were asked to imagine that the distant object was actually placed on their drawing paper and that they could draw along its length. These instructions were designed to ensure that we were dealing with judged or behavioural size and not phenomenal size. The presentation screen was a piece of black card measuring 20×25 in., with a simple elastic attachment for fixing objects at the centre. It stood in a situation which afforded good distance cues, but there were no small familiar objects in its vicinity, other than the test object.

III. RATIONALE OF THE METHOD

Consider the visual size estimate, Y , of a familiar object. This estimate could be based on information from either memory or vision or both, and it is the object of the experiment to find what actually occurs. Suppose that the subjects utilize only visual cues to size. In that event their judgements would be just like judgements of the size of non-familiar objects, where *only* visual information is available. As part of the experiment we therefore obtained judgements, X_2 , from the same subjects of the size of the non-familiar objects, j , k and l . It will be shown that there is a significant variance in X_2 between individuals, and that values of X_2 for the objects j , k and l , correlate significantly together. If the judgements Y of the familiar objects are in fact the same as the X_2 judgements we can thus expect a significant correlation between Y and X_2 . This correlation, r_{Y2} , can therefore serve as an index to the use of vision.

The opposite supposition is that the subjects did not use visual information at all in making the estimate Y , but relied solely on their previous knowledge of the object. In that event they would have made the same judgement even if the object had not been present for them to look at. Judgements, X_1 , based on memory alone, were therefore obtained prior to the visual judgements Y . To the extent that the subjects relied on memory alone, their judgements Y would merely replicate their previous memory judgements, X_1 , and there would be a correlation between Y and X_1 . This correlation, r_{Y1} , can therefore serve as an index to the use of memory.

IV. TREATMENT OF DATA

The first step was to examine the extent and reliability of the individual differences in size estimates. A measure of the over-all self-inconsistency or within-individual error was computed from the duplicated assessments of the individuals in the group, for each type of judgement, Y , X_1 and X_2 , on each object. An estimate of the variance between individuals was derived from the means of the duplicate assessments. The F ratio of the variance estimates between and within individuals then provides a test of the significance of the individual differences. The judgements X_2 of the objects j , k and l , for example, gave F ratios of 5.1, 6.2 and 5.8, respectively, which values are all significant at the 0.001 level (d.f.'s: $n_1 = 37$, $n_2 = 38$). As well as being significant statistically, the differences between individuals were also of a sizeable order: for the above objects, for example, the root mean squares of the variances between individuals were approximately one-quarter of the value of the group's mean judgement of the object, giving, that is, coefficients of variation of the order of 25%. The data for the judgements Y and X_1 on each of the familiar objects were examined in a similar manner. In all cases the variance between individuals greatly exceeded that within, and the differences were statistically significant ($P < 0.001$ except in two of the eighteen cases where $P < 0.01$). The possibility can, therefore, be disregarded that the individual differences in size estimates were merely chance effects arising from the unreliability of the individual assessments.

A central assumption in the rationale of the method was that a significant part at least of the variation in judgements X_2 of the non-familiar objects arose from characteristic individual differences in size constancy, which could thus serve as an index to the use of visual cues to size. Statistical significance is always a relative matter and depends on the size of the residual random error, in this case the self-inconsistency of the subjects' assessments. A possibility was that the consistency of the duplicate assessments was spuriously high, due to each subject merely repeating or copying the judgement he or she happened to make in the first place. Allowance for this in the experimental design was made by separating the duplicate assessments as far as possible in the series of judgements. A further statistical check is possible. If the individual differences in size estimates result from some fairly stable causal factor, then the estimates for different objects will be correlated. Statistically significant correlations between judgements X_2 of objects j , k and l were in fact found, as follows:

$$r_{jk} = 0.64, r_{kj} = 0.82, r_{jl} = 0.60.$$

A further possibility was that these correlations arose in part or entirely from some factor unconnected with the process of estimating size at a distance. The correlation r_{12} , between memory estimates X_1 of familiar size and visual estimates X_2 of non-familiar size, affords a test of this possibility, since there exists no reason other than an underlying general factor why values of X_1 and X_2 should intercorrelate. The three possible values of r_{12} for each familiar object were therefore obtained and found to vary between 0.15 and 0.59 with an average of 0.34, which is significantly greater than zero ($P < 0.05$). An underlying general factor is thus shown to exist: some individuals consistently drew longer lines, others shorter, regardless of the judgement being made. Though this effect will be referred to here as a 'drawing' factor, and could be due to the drawing method itself, it could also be due to a more general motor adjustment or sensory-tonic effect,

or to a still wider tendency to either over- or underestimate in judgement situations generally. Whatever its source, there is evidence that this factor does not account for all the observed correlations; and it can be allowed for statistically. Both X_1 and X_2 can be thought of as predictors of Y , and the multiple correlation $R_{Y.12}$ found of all three variables. If r_{Y1} , say, represents only the drawing factor then, since this source of correlation is also present in the other coefficient r_{Y2} , the addition of X_1 as an additional predictor of Y will result in no improvement in prediction, and $R_{Y.12}$ will not be significantly greater than r_{Y2} . A value of $R_{Y.12}$ significantly greater than r_{Y2} , on the other hand, would indicate that X_1 correlated independently with Y , and that memory factors were operative in the Y judgements. In the same way a value of $R_{Y.12}$ greater than r_{Y1} would indicate that visual factors were operative in the Y judgements. For example, in the results for judgements of a matchstick, shown in the first row of Table 1, r_{Y1} is significantly greater than r_{Y2} , indicating that the influence of memory is predominant. The value of 0.30 for r_{Y2} , however, represents only the drawing factor; there is no evidence that visual cues influenced the Y judgements, since $R_{Y.12}$ is not significantly different from r_{Y1} .

Table 1

Object	r_{Y1}	r_{Y2}	$R_{Y.12}$	$r_{Y1.2}$	$r_{Y2.1}$	Manip. rating	C. Vtn. (X_1) (%)
Matchstick	0.64	0.30	0.65	0.63	—	3.9	18.9
9-inch rule	0.73	0.35	0.74	0.67	—	3.5	15.6
Playing-card	0.59	0.54	0.74**	0.50	0.46	1.5	23.3
Cigarette	0.75	0.47	0.80**	0.66	0.36	2.2	20.0
Nail-file	0.53	0.44	0.63*	0.50	0.35	3.1	26.8
Pocket-comb	0.68	0.44	0.74**	0.50	0.42	2.5	19.8
Medicine-bottle	0.65	0.54	0.74**	0.48	0.33	2.7	29.6
Cigar	0.43	0.75	0.77	—	0.70	0.9	34.0
Penknife	0.36	0.60	0.63	—	0.55	3.3	23.2

r_{Y1} and r_{Y2} are indices to the use of memory and vision, respectively, with the drawing factor not taken into account.

$r_{Y1.2}$ and $r_{Y2.1}$ are indices to the use of memory and vision, respectively, with the drawing factor held constant.

Manip. rating refers to the average rating of manipulative familiarity.

C. Vtn. (X_1) refers to the amount of disagreement in the group's memory standards of familiar size.

* denotes $R_{Y.12} >$ either r_{Y1} or r_{Y2} at the 0.05 level.

** denotes $R_{Y.12} >$ either r_{Y1} or r_{Y2} at the 0.01 level.

As a check, a second method of allowing for the drawing factor was adopted. Utilizing the correlation r_{12} , the partial correlation coefficients $r_{Y1.2}$ and $r_{Y2.1}$ were computed as indicators of memory and visual influences, respectively, with the drawing factor held constant. The two treatments give consistent results, and it will be seen in Table 1 that where there is no significant improvement in the multiple correlation $R_{Y.12}$, then one or other of the partial correlations is not significantly greater than zero. The partial correlations require to be greater than 0.33 for significance at the 0.05 level, and certain values considerably less than this (namely 0.20, 0.10 and 0.23) have been entered in Table 1 as zero. For convenience the results will be discussed in terms of the partial correlation coefficients, rather than the multiple correlations. $r_{Y1.2}$ and $r_{Y2.1}$, like r_{Y1} and r_{Y2} described in section III, are indices of the use of memory and vision respectively, except that in the further partial coefficients the effect of the drawing factor has been partialled out.

V. RESULTS

In Table 1, objects above the horizontal line are 'homogeneous', occurring only or mainly in one standard size; objects below the line are 'heterogeneous', occurring in a variety of sizes.

For the first two objects, the matchstick and the 9-inch rule, the visual coefficients, $r_{12.1}$, were not significantly greater than zero. The subjects apparently utilized only previous knowledge in estimating the size of these objects presented at a distance, though full distance cues were available. However, for the playing-card and the cigarette, also homogeneous objects, the coefficients indicate that both memory and vision influenced the judgements. But the average ratings of manipulative familiarity (sixth column of Table 1) for these objects were significantly lower ($P < 0.05$) than for the former two. Thus the data suggest that the subjects made some use of the available visual information in checking their estimates of the size of playing-cards and cigarettes because, as a group, they were less certain of their knowledge from prior experience with these two objects than with the former two.

For the remaining, heterogeneous objects it was expected that only visual information would be utilized. For the first three of these objects, the nail-file, the pocket-comb and the medicine-bottle, however, there is evidence of both sets of influences at work, with substantial coefficients on the memory side. It would seem that the subjects did in fact have a good idea of the norm for these objects, and that the specimens produced approximated to this norm, which consequently influenced their size estimates.

The last two objects, the cigar and the penknife, are the only ones for which the estimates Y show the exclusive use of vision and no influence of memory. Each of these objects, however, presents, for different reasons, a special case. The cigar has the lowest rating of all for manipulative familiarity, and there is a correspondingly high degree of disagreement in the memory standards of the group for this object, shown in the coefficient of variation of 34% in X_1 (Table 1, seventh column). The subjects apparently had no firmly established subjective standards for the average-sized cigar, and relied on the visual information when asked to estimate the size of an actual cigar at a distance. In the case of the penknife, the effect of a variation in the main procedure of the experiment was tried. As for the other objects, the subjects had initially made estimates from memory of 'the size most commonly met with'. But the penknife presented for visual judgement was much smaller than normal, and an obvious miniature. There was an audible exclamation from the group when this miniature was presented, presumably expressing disappointed expectation. The subjects were nevertheless instructed to estimate the size of the actual penknife presented. In these circumstances the coefficients show no detectable influence on the Y judgements of preconceptions of the average-sized penknife: the subjects merely estimated visually the size of the particular object before them.

A general hypothesis, suggested by and summarizing these findings, would be as follows: when individuals have definite information from memory of the size of distant object, they will utilize this in preference to visual or distance cues to size, even though both sets of information are equally available. Visual cues to size will be utilized to the extent that information from past experience is either lacking or uncertain.

A simple test of this hypothesis was suggested by the fact that the distribution of ratings of the manipulative familiarity of cigarettes was distinctly bimodal. Accordingly, data

were collected from a further group of subjects who made judgements Y and X_1 of cigarettes and judgements X_2 of the squares j and k . The combined data on judgements of squares and cigarettes in this and the main experiment were then dichotomized. Subjects rating cigarettes either two or three for manipulative familiarity were omitted from consideration; those rating zero and unity, on the one hand, or four and five, on the other, were treated as separate groups of 'non-smokers' and 'smokers', respectively. The results, shown in Table 2, are in accordance with the hypothesis: the coefficients for the non-smoker group indicate that vision predominated, though there is evidence of some influence of memory; the visual coefficient is not significantly greater than zero for the smoker group, who apparently relied exclusively on memory.

Table 2. *Judgements of cigarettes*

	$r_{F1.2}$	$r_{F2.1}$	Manip. rating
29 'smokers'	0.67	0.14*	4.3
37 'non-smokers'	0.35	0.63	0.6

* Not significantly greater than zero.

VI. DISCUSSION

A limitation of the results arises from the necessity, in the experimental design, of obtaining the visual judgements (Y) *after* the memory judgements (X_1). This leaves the possibility that the visual judgements (Y) of the familiar objects were influenced by the fact that the subjects had just made non-visual judgements (X_1) of these same objects. It could thus be argued that the weight given to memory in the Y judgements was simply a function of this fact, and was not necessarily the same as the weight which would be given independently, in everyday visual judgements of familiar objects. In other words, the 'past experience' which was effective here could be that introduced in the experiment immediately prior to the visual judgements. But even this limiting case is of considerable interest, since a non-visual, short-term memory factor is shown to be active in a situation, where the visual cues are adequate for, and might be expected to determine, the perception. If immediate prior experience can be active in this way then so may prior experience in the long-term sense.

In their specific implication for size constancy these results support the hypothesis, put forward by Bolles & Bailey (1956), that in familiar situations it is 'not so much the perceived size which is invariant with distance as it is the identification of the object itself'. A wider but more speculative implication is suggested by the fact that the retinal size of a familiar object serves as an indicator of its distance. Thus the Transactionalist group (Kilpatrick, 1952) have demonstrated that when visual cues to distance are reduced or withdrawn, previous knowledge or assumptions concerning the sizes and shapes of things will by themselves determine the perceived spatial relations. The present experiment indicates that, in estimating the size of a distant familiar object, individuals will still rely on their prior knowledge of its size, even to the extent of not utilizing visual cues to distance where these are fully available and adequate to mediate the size estimate. This exclusive use of memory occurred only for the size estimates of the more familiar, homogeneous objects. But familiar environments—as familiar as, say, one's own study or living room—are made up of objects almost all of which have this specific, known size. That most people have readily available, internalized models of familiar scenes is shown

by the ease with which such scenes, and the objects they comprise, can be visualized with the eyes closed. The suggestion is that in familiar surroundings it is the internal rather than the external referent of the retinal image which determines the perceived space; the retinal pattern merely serves to identify the scene, the retinal sizes coupled with known sizes giving the distances of the things within it. Thus perceptual orientation in familiar and non-familiar space may be quite different and to some extent opposite, a 'central' focusing being opposed to a 'distal' focusing of perception. The subjective feeling of familiarity or strangeness in a perceptual situation could also be, in part, a reflexion of this difference in direction of the underlying perceptual process.

The design of the present experiment has certain methodological implications, which may be briefly stated. The investigation required that the separate influences of memory and vision be isolated in a perceptual situation that gave both an equal chance as potential factors; non-interference with the availability, particularly the non-elimination, of either set of cues was essential for the judgements to be analysed. Analysis was made possible by the existence of stable individual differences in each possible condition of estimation, which provided a natural as distinct from an artificially imposed source of variation in the main variables of the problem. In so far as analysis was carried out by means of correlations based on a natural source of variation, the present approach resembles the representative design proposed and developed by Brunswick (1956) as a means of overcoming the limitations of the traditional, systematic design of psychological experiments, with its artificial elimination and tying of variables. Correlations based on a sampling of subjects, rather than objects, appear necessary, however, to separate the memory and visual influences in a situation where both are potential factors. This sampling of subjects is different from—and in some respects the antithesis of—the approach of error statistics, in which variation between individuals is averaged out to give the results for a composite subject. It is different also from the approach of differential psychology, which, when applied to perception, treats the experiment as a kind of test to be related to other tests for the purpose of discriminating between individuals rather than the determinants of the perceptual judgements. One way of viewing the present approach is to conceive the perceptions of different individuals as deriving from their individual 'visual worlds'—which, while largely coincident, are not entirely so; by investigating in the area of non-coincidence, applying correlational techniques, it becomes possible to isolate and study the common determinants of the perceptions. However, the method need not depend on assumptions concerning the source of the individual differences; it is merely necessary to ascertain that stable differences exist for the variables under consideration. Individual differences are in fact a feature of the experimental literature on perception; so the device of deploying them systematically, as in this experiment, may have general applications.

While supporting the principle of a central focusing of perception, and suggesting a new method of investigation, the present experiment alone is insufficient to establish the situational generality of the results, even for size constancy. Only a few objects of the homogeneous and heterogeneous type have been considered, and then only those which are small and manipulable. The sizes of many familiar objects are known in relation to a familiar reference distance rather than by physical contact. The question of the population generality of the results has also to be considered. An average measure of familiarity, for example, is a function of the population making the judgement. Related to this is the entirely unknown relation between the geographical or ecological distribution

of sizes of a specific class of objects, and the kind of distribution of subjective standards to which this gives rise in the population. All these factors stand in need and appear capable of further investigation.

The author wishes to thank Dr B. L. Welch of Leeds University for advice on some statistical points.

REFERENCES

- ALLPORT, F. H. (1955). *Theories of Perception and the Concept of Structure*. London: Chapman and Hall.
- BRUNSWICK, E. (1944). Distal focusing of perception. Size-constancy in a representative sample of situations. *Psychol. Monographs*, No. 254.
- BRUNSWICK, E. (1956). *Perception and the Representative Design of Psychological Experiments*. California: The University Press.
- BOLLES, R. C. & BAILEY, D. E. (1956). Importance of object recognition in size constancy. *J. Exp. Psychol.* 51, 222-5.
- ITTELSON, W. H. (1951). Size as a cue to distance. *Amer. J. Psychol.* 64, 54-67, 188-202.
- KILPATRICK, F. P. (1952). *Human Behavior from the Transactional Point of View*. Institute for Associated Research, Hanover, N.H.
- McKENNELL, A. C. (1953). *Experimental Absolute Scales Compared With a Naturally-Occurring Subjective Standard*. Unpublished M.A. Thesis, University of Manchester.
- McKENNELL, A. C. (1959). Single stimuli judgments based on a standard of familiar size. *Percept. Mot. Skills* 9, 119-26.
- THOULESS, R. H. (1932). Individual differences in phenomenal regression. *Brit. J. Psychol.* 22, 216-41.

(Manuscript received 9 October 1958)

PERIODICITY IN REACTION TIME

By P. H. VENABLES

*Medical Research Council Social Psychiatry Research Unit, Institute of
Psychiatry, Maudsley Hospital, London, S.E. 5*

The distribution of visual reaction times made by six normal and six schizophrenic subjects were found to be multi-modal with peaks in general 0.1 sec. apart. These results are discussed in the light of a cortico-thalamic reverberatory mechanism whose relevance in the light of other psychological experiments is considered.

I. INTRODUCTION

One of the most persistent difficulties in the study of reaction time is that of giving a satisfactory explanation of gross slowness as exhibited for instance in the performance of schizophrenic subjects. During a study (Venables, 1958) of the choice RT of schizophrenic and normal subjects using from one to eight lights a suggestion was made that slow responders were reacting as though there were more alternatives in the situation than were physically present. As the time taken to react to different numbers of alternatives (or number of bits of information, since this study was conceived in information theory terms) was known; the hypothesis was proposed that if a histogram of simple RT's was made it would display modal points at the times taken to process 1, 2, 3, . . . , etc. bits of information. A histogram of all the simple (single choice) RT's of the eight normal and sixteen schizophrenic subjects studied was made and was shown by means of a variogram analysis (Jowett, 1955) to display a periodicity at 0.1 sec. intervals. As the group mean time taken to process one bit of information was 0.094 sec. this result was taken to give some support to the hypothesis. However, because the data analysed were those from a group, the reaction times of whose members might have given rise to apparent periodicity and because individual members of the group had different calculated transmission rates, it was decided to carry out a more extensive experiment and to analyse the results by individual subjects.

II. METHOD

Six schizophrenic and six normal subjects were used. Each subject made 200 or more single choice verbal responses to a light by saying the syllable 'bip', a nonsense syllable which because of its initial plosive sound operated a voice key reliably. The intensity of the stimulus, a diffusely illuminated circular patch 0.75 in. in diameter set in a 2 x 2 ft. matt black board, was 320 foot-lamberts. This intensity was chosen to give optimal RT's with schizophrenics in view of the results showing paradoxical slowing with higher light intensities (Venables & Tizard, 1956). The stimulus light stayed on until switched off by the verbal reaction of the subject which operated a voice key. A fresh stimulus was presented once every 7.5 sec. Reaction times were measured in milliseconds by a 'dekatron' timer. That no source of periodic response could be attributed to the timing circuit or to the reading of the dekatron counters, was checked by passing 480 random times between 0.20 and 0.99 sec. through the apparatus, and reading them at the speed usually adopted under experimental conditions. χ -squared for goodness of fit to equal frequency in each 1/100th second reading between 0.2 and 0.99 sec. was determined. A calculated P of 0.48 indicated no bias in the operation or reading of the timer.

III. RESULTS

A separate histogram on a millisecond scale was plotted for each subject over the range 200-999 msec. Readings above 1 sec. were few and were omitted.

Frequencies of responses grouped by 1/100th sec. for each subject are presented in Table 1.

Table 1. *Frequencies of RT's grouped by 0.01 sec. for normal and schizophrenic subjects*

RT 0.01 sec.	Normals						Schizophrenics						RT 0.01 sec.	Normals						Schizophrenics					
	1	2	3	4	5	6	1	2	3	4	5	6		1	2	3	4	5	6	1	2	3	4	5	6
20	2	1	.	.	.	.	.	.	.	.	.	.	60	.	0	.	0	1	.	7	1	3	3	2	8
21	4	12	.	.	.	.	.	.	.	.	.	.	61	.	0	.	0	0	.	8	4	10	6	9	12
22	38	23	1	.	.	.	.	.	.	.	.	.	62	.	1	.	1	0	.	7	1	8	1	1	5
23	31	5	1	.	.	.	.	.	.	.	.	.	63	.	.	.	0	0	.	0	0	1	0	1	1
24	48	7	2	1	.	.	.	.	.	.	.	.	64	.	.	.	.	1	.	2	0	4	2	7	6
25	10	4	4	0	.	.	.	.	.	.	.	.	65	.	.	.	.	.	.	5	1	8	3	6	7
26	29	14	9	3	.	.	1	.	.	0	.	.	66	.	.	.	.	.	.	2	1	9	3	6	7
27	20	10	19	7	1	.	1	.	.	0	.	.	67	.	.	.	.	.	.	8	0	3	1	4	3
28	2	6	14	8	6	.	0	.	1	0	.	.	68	.	.	.	.	.	.	2	1	4	0	3	3
29	2	8	22	16	4	.	2	.	0	0	.	.	69	.	.	.	.	.	.	6	0	6	1	1	7
30	2	9	10	18	11	.	4	.	1	0	.	.	70	.	.	.	.	.	.	5	0	4	1	1	10
31	2	13	38	21	9	1	8	.	0	1	1	.	71	.	.	.	.	.	.	0	1	7	1	9	11
32	5	5	48	42	32	6	6	.	0	0	0	.	72	.	.	.	.	.	.	1	0	5	0	0	7
33	1	4	3	20	12	5	0	.	0	0	0	.	73	.	.	.	.	.	.	0	0	0	0	1	0
34	0	7	7	20	14	10	5	1	0	0	0	.	74	.	.	.	.	.	.	0	0	4	0	2	6
35	0	7	7	0	2	1	8	2	1	0	1	1	75	.	.	.	.	.	.	0	0	7	1	1	6
36	0	4	3	9	22	10	5	4	2	0	1	2	76	.	.	.	.	.	.	1	0	4	0	1	7
37	1	7	6	9	14	20	3	3	1	0	0	0	77	.	.	.	.	.	.	0	0	1	1	1	4
38	0	10	2	7	16	18	6	7	3	0	0	0	78	.	.	.	.	.	.	0	0	4	1	3	7
39	1	7	1	3	8	17	14	5	2	3	1	0	79	.	.	.	.	.	.	0	0	2	1	2	3
40	0	2	1	3	5	13	7	17	0	1	0	1	80	.	.	.	.	.	.	1	0	0	.	1	7
41	0	6	1	4	13	20	28	27	5	7	4	0	81	.	.	.	.	.	.	1	1	5	.	3	4
42	1	6	1	1	7	32	16	13	3	9	1	0	82	.	.	.	.	.	.	1	0	1	.	1	3
43	0	2	.	1	6	4	1	0	1	0	0	0	83	.	.	.	.	.	.	0	0	0	.	0	0
44	0	3	.	2	1	10	19	14	0	6	2	0	84	.	.	.	.	.	.	0	0	2	.	4	2
45	0	2	.	0	0	3	17	15	0	13	1	0	85	.	.	.	.	.	.	0	0	1	.	0	4
46	0	2	.	1	4	3	18	8	0	17	2	0	86	.	.	.	.	.	.	0	0	0	.	0	1
47	0	3	.	0	3	6	17	10	2	12	4	0	87	.	.	.	.	.	.	0	1	2	.	2	4
48	0	1	.	1	2	4	10	10	4	11	3	2	88	.	.	.	.	.	.	0	.	0	.	1	1
49	0	3	.	0	1	6	12	7	2	15	1	0	89	.	.	.	.	.	.	0	.	0	.	0	3
50	0	1	.	0	1	1	13	7	4	12	4	6	90	.	.	.	.	.	.	0	.	0	.	1	5
51	0	2	.	0	1	2	29	13	5	31	10	5	91	.	.	.	.	.	.	2	.	2	.	1	2
52	1	2	.	1	2	5	18	4	5	12	8	1	92	.	.	.	.	.	.	.	.	.	.	0	1
53	.	0	.	1	0	0	0	1	1	1	0	0	93	.	.	.	.	.	.	.	.	.	.	0	1
54	.	0	.	0	0	2	10	3	5	5	7	2	94	.	.	.	.	.	.	.	.	.	.	1	5
55	.	0	.	0	0	0	8	4	6	8	6	1	95	.	.	.	.	.	.	.	.	.	.	0	3
56	.	0	.	0	0	0	8	3	4	7	2	5	96	.	.	.	.	.	.	.	.	.	.	0	1
57	.	0	.	0	0	0	4	3	2	3	2	6	97	.	.	.	.	.	.	.	.	.	.	0	5
58	.	0	.	0	0	1	7	3	0	2	5	7	98	.	.	.	.	.	.	.	.	.	.	3	0
59	.	.	.	0	0	.	9	3	1	4	6	12	99	.	.	.	.	.	.	.	.	.	.	0	2

From these results it is seen that for the normal subjects there appear to be modal points every 0.1 sec. at 0.22, 0.32 and 0.42 sec. with occasional sporadic responses at 0.52 and 0.62 sec. There also appears to be evidence of subsidiary modal points at around 0.26-0.27, 0.36-0.37 sec.

In the case of the schizophrenic data modal points are again found at 0.1 sec. intervals at 0.41, 0.51, 0.61, 0.71 sec. with occasional grouping of responses at 0.31 and 0.81 sec. There does not in the schizophrenic data appear to be any evidence for secondary modal points at 0.06 or 0.07 past each 0.1 sec. interval.

To clarify the position the frequencies in Table 1 were summed over subjects separately for normals and for schizophrenics and then the sums of frequencies of RT's of values $0.21 + 0.31 + 0.41 + \dots$, etc.; $0.22 + 0.32 + 0.42 + 0.52$, etc.; $0.23 + 0.33 + 0.43$, etc. taken. By these means random variations ought to cancel out and coincident variations show up more clearly. The result of these summations is shown in Fig. 1, where it is apparent there are main modal peaks at 0.01 and 0.02 sec. for schizophrenics and normals respectively. The points from 0.06 to 0.09 sec. have a virtually equal frequency distribution.

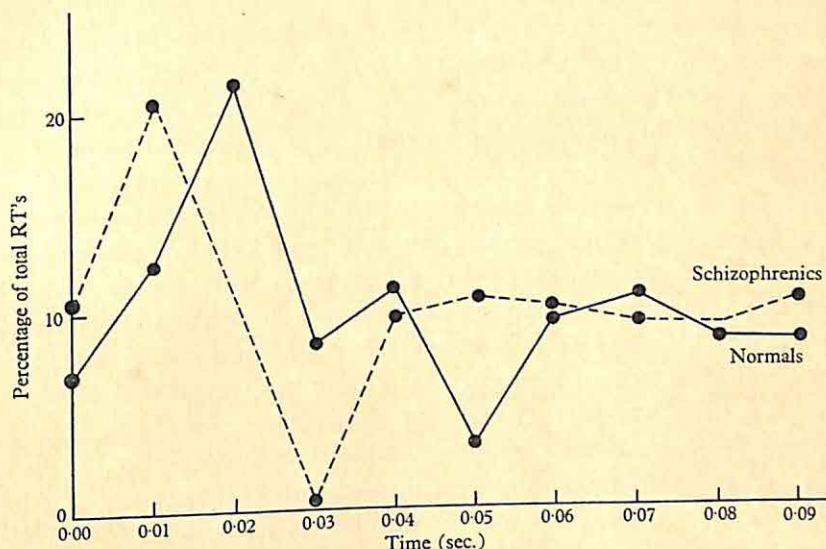


Fig. 1. Percentages of RT's falling in superimposed and cumulated cycles of 0.1 sec. (see text).

There are points at 0.03 sec. for schizophrenics and normals and 0.05 sec. for normals where relatively few responses are made. These results are shown by the use of χ -squared to differ from chance at well beyond the 0.001 level of confidence.

It thus appears that slowness in both normals and schizophrenics is not just random but is in a large number of cases as though having missed one cycle of excitation the response is elicited by one of a number of subsequent cycles appearing at a frequency of 10/sec.

That this periodicity has apparently no relation with the time taken to process information is indicated by the same periodicity appearing in all subjects whose calculated times per bit on the type of RT task here employed range from 0.056 to 0.104 sec. in the case of the schizophrenics and from 0.064 to 0.114 sec. in the case of normal subjects.

IV. DISCUSSION

Any attempt to relate data of this nature to neurophysiological phenomena must remain speculative. However, it may be possible to make some decision between various hypotheses which may usefully delimit the field of future investigation.

In general two facts have to be explained, the source of the apparent periodicity of RT, and the slowness of RT due to delay of response to later cycles of periodicity. It is doubtful how far it may be useful to speculate on the 0.01 sec. difference between the modal peaks in schizophrenic and normal subjects.

As far as sources of periodicity are concerned there would seem to be two possibilities. One where the periodicity is related in time to the onset of the stimulus, and the other where the stimulus arrives in random relation to an already on-going source of periodicity.

As it should logically be impossible to detect any periodicity in stimulus-response time due to a signal arriving in the cortex in random relation to an on-going central cyclical mechanism and in view of the conflicting results of Lansing (1957), O'Hare (1954) and Walsh (1952) in attempting to relate RT to the α -rhythm the latter possibility will not be considered. Chang (1950), following previous work, for example, by Bartley & Bishop (1933), Adrian (1941), Bremer (1949) and Marshall, Talbot & Ades (1942), has described cortico-thalamic reverberatory mechanisms actuated by afferent stimulation from visual, auditory and somesthetic sources. 'In response to an afferent stimulation the sensory cortex gives rise to a potential change consisting of a primary response followed by a series of from 3-13 regularly spaced waves.' The frequency of response is independent of intensity of stimulation. In the case of the visual and the somesthetic systems there are successions of waves following the primary response at intervals of about 100 msec. In the case of the auditory system secondary waves appear at about 80 msec. intervals, (Rosenzweig & Rosenblith (1953), however, give evidence for secondary responses at 100 msec. intervals in the auditory cortex). 'Responses of a pattern similar to that recorded from cortical structures can be recorded also from corresponding thalamic structures. Such responses can be completely abolished, however, either by the removal of the concerned cortex or by the interruption of cortico-thalamic connexions.' Chang also states that in view of further experimental data it is implied 'that the thalamo-cortical pathway for the repetitive discharges must be independent of that occupied by primary afferent volleys'.

We here have details of a neurophysiological mechanism which may be a basis of explanation of the RT findings. From this point on we can do no more than speculate as to the possible time relationships involved as components of the total RT.

Monnier (1952), in a study of response to photic stimulation using E.R.G., E.E.G. and E.M.G. techniques gives various figures which may serve as starting-points. As the time taken for the first response to appear in the cortex subsequent to a weak photic stimulus he gives a mean figure of 65 msec. Monnier then deduces a value of 75-80 msec. as a figure for what he calls the opto-motor integration time. This is the time taken by central mechanisms between the optic and the motor cortex. For the time from the activation of the motor cortex to the response, in the present experiment, being recorded as a sound we have to rely a great deal on guess work. If we were using a hand reaction we could expect about 10-15 msec. to be concerned in nerve conduction, e.g. Wagman & Lisse (1952) and 30-40 msec. to be taken up by time elapsing before action potential is translated into movement (Vorckel, 1927 quoted by Woodworth, 1938). This gives an average figure of 47 msec. spent on the motor side of a manual RT response. Cattell (1886) quoted by Woodworth (1938) gives a figure of 27 msec. for the difference between manual and verbal response to a light. We can thus guess at about 75 msec. as the time taken on the motor side of a verbal response. The total time to be expected over a visual-verbal RT is thus in the region of 215 msec. which accords very well with lowest RT recorded in the present study.

Now if the reaction is not made in response to the primary sensory input we can expect

a secondary peak of potential in the optic cortex from the cortic-thalamic mechanism after 100 msec. which will give rise to a response at 315 msec., and so on. A paradox might at first appear in the light of Chang's experimental data, as he finds a more or less exponential decrease in the size of successive reverberatory waves, which might make it seem unlikely that a motor mechanism which was not fired by a large primary input would be fired by a smaller secondary or later one. The operation of the mechanism of temporal summation would nevertheless adequately provide an explanation. Successive 0.1 sec. potential would summate until a value which would activate later neurones is reached. In addition, in a later article, Chang (1952) shows that 'cortical response to the stimulation of the lateral geniculate body can be potentiated by continuous illumination of the retina as manifested by the lowering of the threshold and increase in size of response'. This potentiation could apparently take place in the presence of very small photic stimulation. Light transmitted through closed eyes was sufficient to cause potentiation and complete blindfolding was necessary to modify the effect.

There is an increased size of secondary response up to about 5 sec. and even if the original stimulus was only a single flash it is likely that late responses will take place unless there is complete darkness following the flash. Even then it would seem, according to Chang, that there can be some potentiation by input from other modalities all of which it is presumably impossible to eliminate.

It would appear that in the work of Chang we have a description of a neurophysiological mechanism which could be used to explain the present findings. It must be remembered that Chang's work is mainly on cats and his findings need not be transferable *in toto* as an explanation of human behaviour. There is, however, sufficient in his work to give rise to testable hypotheses in the psychological field. One directly testable is that the cortico-thalamic reverberation in the auditory modality taking place at a rate of 0.08 sec. per cycle should show up in periodicity in the auditory RT. In the analysis of some data gathered from the performance of thirty-two schizophrenics on a RT task with manual response to a stimulus of 1000 cyc./sec. and intensity 70 d.b., a variogram analysis of the distribution of single RT's showed periodicity at 0.08 and 0.10 sec. There seemed to be some evidence in favour of the former figure on examination of the distribution of RT's. This is, however, slight evidence in support of the hypothesis as it is based on group data with a small number of RT's from each subject, and the variogram analysis derived from this data did not give a sufficiently clear indication of periodicity to be very decisive.

One aspect of the data which does not appear at first to fit directly with the possible neurophysiological mechanism is the shape of the distribution of RT's. We might expect the responses of any individual clearly grouped round a single point with a distribution about 15-20 msec. on either side as far as can be judged from the figures given by Monnier (1952), for the standard deviations of the elements of reaction time. Instead of this we have some 50 msec. containing a number of responses before this expected peak of maximum number of responses followed by 10-20 msec. of comparative absence of response. While the latter unresponsive period is in accordance with expectation, it is the 50 msec. before the peak that requires explanation. What seems most likely is that in a mechanism which, if temporal summation is operative, depends on a process of integration, potentials will reach a threshold value at some time on the rising phase of a pulse while the subsequent falling phase will have no effect on the level reached.

A well known phenomenon which gives confirmatory value to the postulated explanatory mechanism is that of brightness enhancement (the Brücke effect), of Bartley (1938). In this, if a series of flashes at a rate of from 8–10 flashes/sec. is presented in the visual field the intensity of the flashes appears subjectively greater than constant illumination at the same intensity. While this phenomenon has been related to the frequency of the α -rhythm it necessitates a mechanism whereby the internal phases of excitation are related in time to the onset of the initial flash. The mechanism outlined by Chang would seem to provide this necessary function, whereas unless it is postulated that the α -rhythm 'locks on' to the flash rate it is difficult to see how this would account for the enhancement effect, as has been suggested. Also consonant with the postulated reverberatory mechanism are the results of experiments by Cheatham & White (1952) and (1954), who found that auditory or visual pulses presented at a rate higher than 10 cyc./sec. were not all perceived. For example, auditory pulses presented at 30 cyc./sec. were perceived only at a perceptual rate of 10 cyc./sec. Presumably pulses falling in the time between the elicitation of a primary response and the appearance of secondary ones fall in a refractory state of the mechanism which does not allow these pulses to be registered.

V. CONCLUSIONS

The cortico-thalamic reverberation system would appear to offer a suitable explanatory mechanism for the findings presented. It is productive of further testable experiments and seems in line with other notions about schizophrenic mechanisms, namely the state of inhibition of the cortex which might prevent the excitation of the motor system by the primary sensory input; and also with experiments, for example Tizard & Venables (1957), concerning the effect of additional sensory input upon speed of RT.

One interesting query is why this periodicity has not been noticed by other workers in the well-trodden field of RT. Several answers are possible. One, that until recently there have not been timing mechanisms available which would make possible the measuring and reading of milli and hundredths of seconds with the same ease and accuracy as larger values. Further, workers in the normal field have largely been interested in obtaining stable figures from highly practised subjects and have rejected what is in the present instance the earlier more variable results; there is a tendency for later responses in a series to fall on only one or two modal points. Thirdly, workers in the field of schizophrenia have been interested in showing gross mean difference or greater variability between schizophrenic and normal subjects and have not plotted large numbers of individual times.

Finally, it may be objected that the multitudes of studies of reaction times carried out in the last century do not in fact show differences due to biological and experimental variables as gross or as regular as 100 msec. in the case of the visual modality or 80 msec. in the auditory modality. This does not seem necessary to invalidate the present position insofar as different variables can effect the components of reaction time to different extents. For example, age may effect the conductance time of the motor impulse for a manual reaction by as much as 10 msec. Dealing with complexity occasioned by the transmission of one bit of information may effect the time taken in presumably some central mechanism such as the optomotor integration time of Monnier by from about 0.06 to 0.15 sec. The time at which this stimulus is to appear can involve 18 msec. per

bit of stimulus uncertainty (Klemmer, 1957). These examples show that to study the components of the total reaction time we must be extremely careful in holding constant those variables known to affect different elements of the total RT.

After this paper was prepared, further evidence of 100 msec. periodicity has become available. Augenstine (1958) states that 'analysis of response times indicated that the performance of (these) simple tasks is undoubtedly quantized in units of 100 msecs.'

REFERENCES

- ADRIAN, E. D. (1941). Afferent discharges to the cerebral cortex from peripheral sense organs. *J. Physiol.* **100**, 159-91.
- AUGENSTINE, L. G. (1958). Human performance in information transmission. Report R75, Part 4, Control Systems Laboratory, University of Illinois.
- BARTLEY, S. H. & BISHOP, G. H. (1933). The cortical response to stimulation of the optic nerve of the rabbit. *Amer. J. Physiol.* **103**, 159-72.
- BARTLEY, S. H. (1938). Subjective brightness in relation to flash rate and light-dark ratio. *J. Exp. Psychol.* **23**, 313-19.
- BREMER, F. (1949). Consideration sur l'origine et la nature des 'ondes' cerebrales. *E.E.G. Clin. Neurophysiol.* **1**, 177-93.
- CHANG, H.-T. (1950). The repetitive discharges of cortico-thalamic reverberating circuit. *J. Neurophysiol.* **13**, 235-57.
- CHANG, H.-T. (1952). Cortical response to the stimulation of the lateral geniculate body and potentiation thereof by continuous illumination of the retina. *J. Neurophysiol.* **15**, 5-26.
- CHEATHAM, P. G. & WHITE, C. T. (1952). Temporal numerosity: I. Perceived number as a function of flash number and rate. *J. Exp. Psychol.* **44**, 447-51.
- CHEATHAM, P. G. & WHITE, C. T. (1954). Temporal numerosity: Auditory perception of number. *J. Exp. Psychol.* **47**, 425-8.
- JOWETT, G. H. (1955). The comparison of measures of industrial time series. *Appl. Statistics*, **4**, 32-46.
- KLEMMER, E. T. (1957). Simple reaction time as a function of time uncertainty. *J. Exp. Psychol.* **54**, 195-200.
- LANSING, R. W. (1957). Relation of brain and tremor rhythms to visual reaction time. *E.E.G. Clin. Neurophysiol.* **9**, 497-504.
- MARSHALL, W. H., TALBOT, S. A. & ADES, H. W. (1942). Cortical response of the anesthetized cat to gross photic and electrical afferent stimulation. *J. Neurophysiol.* **6**, 1-15.
- MONNIER, M. (1952). Retinal, cortical and motor responses to photic stimulation in man. *J. Neurophysiol.* **15**, 469-86.
- O'HARE, J. J. (1954). The variability of auditory and visual reaction time with change in amplitude and phase of alpha rhythm. *Amer. Psychologist*, **9**, 444.
- ROSENZWEIG, M. R. & ROSENBLITH, W. A. (1953). Responses to successive auditory stimuli at the cochlea and at the auditory cortex. *Psychol. Monog.* **67**, No. 13.
- TIZARD, J. & VENABLES, P. H. (1957). The influence of an extraneous stimulation on the reaction time of schizophrenics. *Brit. J. Psychol.* **48**, 299-305.
- TRAVIS, L. E. (1929). The relation of voluntary movement to tremors. *J. Exp. Psychol.* **12**, 515-24.
- VENABLES, P. H. (1958). Stimulus complexity as a determinant of the reaction time of schizophrenics. *Canad. J. Psychol.* **12**, 187-90.
- VENABLES, P. H. & TIZARD, J. (1956). Paradoxical effects in the reaction time of schizophrenics. *J. Abnorm. (Soc.) Psychol.* **53**, 220-4.
- WAGMAN, I. H. & LESSE, H. (1952). Maximum conduction velocities of motor fibres of ulnar nerve in human subjects of various ages and sizes. *J. Neurophysiol.* **15**, 235-44.
- WALSH, E. G. (1952). Visual reaction time and the alpha rhythm, an investigation of the scanning hypothesis. *J. Physiol.* **118**, 500-8.
- WOODWORTH, R. S. (1938). *Experimental Psychology*. New York: Holt.

SERIAL ORDER INTRUSIONS IN IMMEDIATE MEMORY

By R. CONRAD

Medical Research Council, Applied Psychology Research Unit, Cambridge

The theory that serial order intrusions in immediate memory lead to recall errors was tested experimentally. It was found that increasing the time interval between successive 8-digit messages had no effect on recall performance. Analysis of the response errors showed that with a short interval between messages, serial order intrusions occurred, but they did not occur with long intervals. Thus the amount of error appeared to be independent of the number of intrusions. It is suggested that intrusions occur after recall has failed—there being no causal relationship.

I. INTRODUCTION

A recent paper by the author (1959), concerned with immediate memory for digits, reported a curious phenomenon noted by Melton & Von Lackum (1941) for nonsense syllables. It appeared that when a digit was forgotten, there was a tendency for subjects to replace it by the digit which occurred in the same serial position but in the previous message. This seemed to be a highly specific example of intrusions from previous lists which have been extensively reported in paired associate learning studies, and in studies of immediate memory for verbal material (Werner, 1947).

The essential condition for the phenomenon is that shortly after the subject has repeated back one message, which he has heard only once, another is presented for immediate recall. Any errors which then occur in the second message may take the form of serial order intrusions from the previous one. In this case one is forced to consider the relationship between the intrusion and forgetting. Traditionally it has been assumed that the subject has forgotten because of the intrusion, i.e. intrusions represent a form of response competition as suggested by Morgan & Underwood (1950) which results in inhibition. This view would link intrusions with proactive inhibition which Underwood (1949) has defined as 'the decrement in recall of a task as a result of prior learning of other material'. In other words, intrusions proactively inhibit correct recall and lead to errors.

One of the difficulties entailed in this argument is that unless one can assert that the effects of proactive inhibition cease at the next message, in a test in which a subject hears and recalls a number of messages in succession, recall ought to become progressively worse throughout the test. Underwood (1957) indeed has collated some striking material on paired-associate learning, showing that list-learning in the laboratory is negatively correlated with the number of lists learned in previous tests. But deterioration of this kind has never been reported in the simple case of immediate recall of short lists. It may therefore be that serial order intrusions do not directly lead to forgetting. In this case the postulate is that having forgotten, the subject 'selects' the intruded digit from a set of alternative possibilities.

One could compare these two views quite simply by finding a test condition which did not show serial order intrusions. If the response competition theory is correct, recall should be better in this condition than in a condition in which intrusions occurred. However, more information would be obtained from an experiment by postulating that serial order intrusions are a function of the time interval between successive messages;

that the longer the interval, the fewer would be the intrusions. The relationship in a situation of this kind, between proactive inhibition and interval between successive messages, has not been described in the literature, but one must on commonsense grounds expect that the effect would diminish with time.

On a proactive inhibition/response competition theory, one would expect that long intervals between messages would result in fewer serial order intrusions, and thus fewer recall errors. On an alternative view that forgetting is independent of intrusions, one would on decay theory expect long intervals between messages to lead to fewer intrusions, but to no change in the number of recall errors.

II. PROCEDURE

The subjects, who were twenty-four naval ratings, were each tested under three conditions, which differed only in the interval provided between the successive messages. During any one condition the interval was the same, and was either 15, 25 or 40 sec.

Ferrograph tape-recordings were made of 8-digit messages, which were spoken at a rate of 2 digits per sec. in a fairly regular monotone. Eight successive messages had the same interval between the end of one and the beginning of the next; the next eight messages were separated by a different interval and so on, for the forty-eight messages which comprised each tape. Three tapes were prepared, the same messages being used in the same order on each tape. The tapes were distinguished only by the order in which the critical time intervals occurred:

Tape 1: 15, 25, 40, 40, 25, 15 sec.

Tape 2: 40, 15, 25, 25, 15, 40 sec.

Tape 3: 25, 40, 15, 15, 40, 25 sec.

In this way, the three test conditions were balanced for order of presentation, when the subjects were assigned in sequence to one of the three tapes, eight subjects to each tape.

The procedure for each subject was as follows. He was told that he would hear an 8-digit number, and that when he had heard it, he should repeat back what he thought it was. This response was recorded by E. After an interval of 15 (25, 40) sec. he would hear another number and so on. After eight messages had been disposed of, the tape was stopped and O was told that for the next eight messages there would be an interval between each of 25 (15, 40) sec. This procedure was continued throughout the session, which lasted for about 25 min.

The individual messages themselves were made up from tables of random numbers with certain constraints. No digit occurred twice in a message and runs (234, 876, etc.) of more than two digits were disallowed. No digit appeared in the same serial position in consecutive messages.

III. RESULTS

The frequency of occurrence of serial order intrusions from the previous message was examined first. An intrusion was counted when a digit was wrong, and was replaced by the digit which was given for the same serial position in the previous message, whether this latter digit was right or wrong, e.g.:

	Stimulus	Response
Message 1	48392716	48392716
Message 2	81470639	81472639
Message 1	07158294	07158264
Message 2	96372580	96372560

Errors which were not intrusions were also counted. Since reversal (transposition) errors were found to occur with approximately equal frequency in each condition, they were included when searching for intrusions. Errors of omission were excluded from the analysis because they could not be classified, occurred equally frequently in all conditions, and were fewer than 10% of all errors. One would expect, by chance, that one-ninth of the errors would appear as intrusions.

Table 1 shows, for the three conditions, the relative frequency of intrusion and non-intrusion errors, together with the values which would be expected by chance. This result shows that the occurrence of serial order intrusion depends on the interval between successive messages. By the time this interval was 40 sec., there were no intrusions.

Table 1. *Analysis of errors*

Interval	15 sec.		25 sec.		40 sec.	
	Intrusion	Other	Intrusion	Other	Intrusion	Other
Observed	70	272	48	308	40	322
Expected	38	304	40	316	40	322
χ -squared	11.12		0.67		0	

Table 2. *Number of messages wrongly recalled*

Interval	15 sec.	25 sec.	40 sec.
Number wrong (out of 384)	154	158	159

Table 2 shows, for the twenty-four subjects, the total number of messages wrongly recalled. It will be immediately obvious that this value is independent of the time interval between messages.

Finally an examination of the results for progressive effects of proactive inhibition fails completely. Towards the end of the list of forty-eight messages, recall is neither better nor worse than at the beginning.

IV. DISCUSSION

The occurrence of serial-order intrusions first reported by Melton and Von Lackum (1941) is again confirmed. These authors clearly associate the phenomenon with inhibition and conclude that the competition between the intrusive and the correct response leads to blocking and hence recall decrement. This is the view which one might regard as the conventional one, though indeed no critical evidence has been advanced to show that intrusions are responsible for recall decrement.

The present evidence suggests that in fact this is not the case. Table 1 in conjunction with Table 2 both show that recall errors are quantitatively independent of intrusions. To be exact, it could be argued that intrusions do create errors, that the fact that there are no fewer errors with long intervals between messages, is because some other unspecified factor, not present with short intervals, exactly counter-balances the absence of intrusions. On the face of it this seems an unlikely explanation, and one might prefer the view that, since the absence of intrusions does not diminish error, and the presence of them does not increase error, then intrusions cannot be competing with the correct response.

Equally one can say that intrusions are not an expression of proactive inhibition, since the presence of proactive inhibition can only be detected by recall decrement (Under-

wood, 1949). Furthermore it seems unlikely that proactive inhibition occurs at all in this kind of situation where a sequence of different messages are alternately presented and recalled. If it did occur, it would surely be weaker in the long interval condition, when we would expect to find fewer errors, but do not. Again if Underwood's (1957) data generalized to this situation, so that the more messages one learned the stronger the effects of proactive inhibition became, messages would become increasingly difficult to learn. In the present experiment there is no evidence at all for this. Indeed, if the principle is generally valid, one wonders how anything is ever learned.

The results of this experiment more readily suggest the conclusion that forgetting has occurred before the intruding digit intrudes—rather than because it has. It is not a competing stimulus. On the simplest decay theory of immediate memory, increasing the time interval between messages would reduce the likelihood of an intrusion. The only 'competition' that would occur would be between the digit determined by serial position, and the other digits. With a 40-sec. interval between messages, all digits are equally likely to be selected.

It is an interesting thought that the subject is usually aware of a gap in the record of his recall. Failing to recall the required 8-digit message, he does not make a 7-digit response. Nor does he give seven correct digits and guess at the eighth last one. The order information in the message is, as Broadbent (1957) and Brown (1958) have suggested, of importance. What probably happens is that, expectedly, order information decays independently of item information. Thus fewer serial order intrusions occur after a long interval since the previous message. When an item is forgotten, if the order information in the previous message is sufficiently preserved it may well affect the choice of a replacement.

REFERENCES

- BROADBENT, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychol. Rev.* **64**, 205-15.
- BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quart. J. Exp. Psychol.* **10**, 12-21.
- CONRAD, R. (1959). Errors of immediate memory. *Brit. J. Psychol.* (In the Press.)
- MELTON, A. W. & VON LACKUM, W. J. (1941). Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.* **54**, 157-73.
- MORGAN, R. L. & UNDERWOOD, B. J. (1950). Proactive inhibition as a function of response similarity. *J. Exp. Psychol.* **40**, 592-603.
- UNDERWOOD, B. J. (1949). Proactive inhibition as a function of time and degree of prior learning. *J. Exp. Psychol.* **39**, 24-34.
- UNDERWOOD, B. J. (1957). Interference and forgetting. *Psychol. Rev.* **64**, 49-60.
- WERNER, H. (1947). The effect of boundary strength in interference and retention. *Amer. J. Psychol.* **60**, 598-607.

(Manuscript received 15 January 1959)

STIMULUS-RESPONSE THEORY AND EXPECTANCY

By MICHEL TREISMAN

Institute of Experimental Psychology, University of Oxford

In the debate between 'expectancy' and '*S-R*' theory the latter has appeared to be objective and precise and the former vague. Since fairly detailed models have now been put forward for expectancy theory, the adequacy of '*S-R*' answers to 'expectancy' questions deserves re-examination. These answers are found to depend mainly on the employment of 'secondary reinforcement' and the 'fractional anticipatory goal response'. The explanations of such problems as latent extinction, reasoning and drive discrimination, which these afford, are shown to be fallacious.

I. INTRODUCTION

The concept of the stimulus-response (*S-R*) link implies that reception of a particular stimulus can be physiologically associated with the production of a particular reaction. It has its roots in associationism, conditioning theory and the knowledge about peripheral reflexes which was available to the early behaviourists. To them it appeared to be the only possible basis for a rigorous, determinate explanation of behaviour, the only alternative to the acceptance of vitalism and vague anthropomorphic principles such as 'instinct' or 'purpose'. From this has developed the programme of the neobehaviourists, represented particularly by Hull and his followers, which is basically the explanation of all behaviour as the establishment of particular assemblages of *S-R* links. A reinforcement principle, derived from Thorndike's Law of Effect, and other elaborations, has been added (Watson, 1924; Sherrington, 1906; Pavlov, 1927; Hull, 1943, 1952; Spence, 1951*a, b*, 1956).

Despite its successes, *S-R* theory met with difficulties in the explanation of behaviour showing 'choice' or 'purpose'. On the other hand, development in fields such as cybernetics and the study of servomechanisms has led to more sophisticated concepts of possible modes of central nervous organization, and models have been put forward which explain the patterning of behaviour in other terms than *S-R* contingencies. Examples are Tinbergen's (1951) account of directiveness in instinctive behaviour, and Deutsch's (1956*a*) explanation of goal-seeking in learned behaviour. In the latter scheme a distinctive external stimulus will feed into a corresponding central mechanism called a 'link'. If an external stimulus is present to the animal, and the corresponding link is receiving excitation from a drive centre directly or through other links, it will produce motor activity. Feedback from the stimulus will so control this activity as to cause the animal to approach it. Learning a maze will consist in forming a chain of such links in the order of the maze stimuli, with the first link connected to a drive centre. Since changes near the goal affect the passage of excitation down the chain, this scheme accounts economically for learning or extinction without performance or reward, and other 'expectancy' phenomena.

Since we now have alternative accounts of expectancy and related phenomena, it is of interest to determine their adequacy. The present paper is an examination of some of the better worked-out applications of *S-R* theory to the more important phenomena in this area, such as drive discrimination, incentive motivation, reasoning, latent

learning and extinction. Concepts frequently employed for this purpose, such as 'secondary reinforcement' and the 'fractional anticipatory goal response' will be critically examined, in so far as this is relevant to the purpose of the paper.

II. THE FRACTIONAL ANTICIPATORY GOAL RESPONSE (r_G)

This was first proposed by Hull in 1931. He suggested that of the consummatory response to the goal (R_G), that fraction which was compatible with other ongoing responses might be evoked by stimuli preceding the goal as a covert response (r_G), and that the proprioceptive stimulation consequent on this response (s_G) might serve to evoke further responses. There is little observational evidence for r_G 's. Conditioned responses resembling consummatory responses do occur (Pavlov, 1927). But it is doubtful if they can be considered fractional; when adequately described they often appear qualitatively different from the consummatory response (Zener, 1937; Hilgard & Marquis, 1940). Miller (1935) claimed to have observed fractional responses in a maze-running situation. He found that rats tended to run on the right side of a runway when the reward lay to the right at the end of it, and on the left when a left turn was required. What particular fraction of the final left or right turn, with its characteristic proprioceptive stimulation, might be contained in running forward on the left or right, was not discussed. This behaviour seems much better considered as an instance of the tendency of rats to prefer paths in the direction of their goal, as described by Tolman (1932).

Although Hull (1931, 1952) initially had the consummatory response in mind as the source of the r_G , he later assumed they could derive from responses to distinctive features of goal boxes (Hull, 1935), and Ritchie (1959) has postulated distinctive r_G 's from eating responses made in differently coloured goal boxes.

Although they have not been demonstrated, I shall assume that r_G 's might occur, and examine how S - R theory, enriched by this development, handles expectancy.

III. SECONDARY REINFORCEMENT

It has often been noted that learning may be produced or extinction retarded by stimuli which have been associated with a reward. For Tolman (1932) this is an expression of sign-Gestalt-expectations, for Deutsch (1956*a*) the result of a connexion between the link corresponding to the stimulus, and a drive centre. S - R theory treats it as an extension of the reinforcing effect of a primary reward.

In *Principles of Behavior* (1943) Hull discussed the problem: the direct effect of primary reinforcement in strengthening habit connexions only extends back some 20 sec. Yet persistent habit chains of much greater length are easily established. How can they be accounted for?

Basic to Hull's learning theory is the assumption that habits are built up by the occurrence of a stimulus and a response in close association, followed (with a delay of not more than 20 sec.) by a reinforcing state of affairs. No limitations are placed on the nature of stimulus, response or reinforcing state of affairs: any stimulus can be linked to any response by any coincident state of affairs which has reinforcing power. And a state of affairs which is reinforcing is equally reinforcing whatever the stimuli and response with which it happens to coincide.

He extended this idea to cover secondary reinforcement: any neutral stimulus frequently coinciding with a reinforcing state of affairs would itself acquire reinforcing properties. Further he suggested that these might derive from a fractional component of the goal reaction becoming conditioned to the stimulus. Consider his example: a rat learns a habit sequence—say, running a maze—that brings it to a reinforcing state of affairs, G . The stimuli (choice-points) at which it has to make the appropriate responses are numbered from S_1 (the start) to S_5 . To each stimulus the appropriate response R_1 to R_5 , must be made, to bring the rat to G , the reinforcing state of affairs. How has this habit chain been established? The final stimulus, S_5 , has had the final response (R_5), preceding G , conditioned to it by the reinforcing effect of G . But how have the earlier habits, $S_1 \rightarrow R_1$, $S_2 \rightarrow R_2$, etc., been established? They are out of the reach of the primary reinforcement provided by G . Hull's answer is that when G was reinforcing the connexion $S_5 \rightarrow R_5$ it was also giving S_5 secondary reinforcing power. By virtue of this S_5 could then reinforce a connection between S_4 and R_4 and make R_4 secondarily reinforcing, and so on. Attributing the secondary reinforcement to a fractional component of the goal reaction, he states (p. 100) that the fractional component 'as both a conditionable process and a reinforcing agent, would be passed back through the sequence', being conditioned successively to $S_5, S_4, \dots, S_1$. The earlier habit segments in the sequence would be built up by secondary reinforcement alone.

This appears a satisfactory and ingenious solution of a difficult problem until we consider what will happen during extinction. Hull assumes that when G is removed the weaker reinforcing power of the fractional component of the goal response will not be able to sustain the connexion $S_5 \rightarrow R_5$, and this will weaken, as will also the secondary reinforcing power of S_5 , so that the habit sequence will gradually collapse, $S_4 \rightarrow R_4$ becoming weaker, and then $S_3 \rightarrow R_3$, and so on. But this does not accord with the previous assumptions. If secondary reinforcement was adequate to build up and maintain the earlier segments of the habit chain, it should be quite adequate to continue to maintain them, and to maintain the later segments. And there is no reason why it itself should be reduced: if the secondary reinforcement from a fractional goal response (e.g. that evoked by S_3) was adequate to initiate, strengthen and maintain a connexion between an earlier stimulus (S_2) and the fractional component of the goal reaction, so as to make S_2 secondarily reinforcing, then the fractional goal responses which the various goal-box stimuli will evoke (in an irregular order resulting from the more-or-less random activity of the animal in the goal-box) will strengthen each other, each evocation of the fractional goal response strengthening any recent stimulus-fractional component associations, so that the fact that G no longer occurs provides no basis for a reduction in established secondary reinforcement, and the complete disintegration of the habit sequence cannot be accounted for.

Hull comes near appreciating this point when he supposes that secondary reinforcement need not necessarily extinguish if primary reinforcement has been 'profound' and so set up functionally autonomous habits. But he does not realize that, on the scheme presented, if secondary reinforcement is adequate to establish learning it is adequate to prevent extinction in all cases, and not only when the habit is grossly overlearned. This is a necessary consequence of the principle that secondary reinforcement derives from frequent coincidence of a stimulus and

a reinforcing state of affairs, and is still present if the secondary reinforcement is attributed to the occurrence of r_G 's.

The argument can be taken further if the secondary reinforcement is attributed to s_G , the interoceptive stimulus which is the physiological consequence of the anticipatory goal reaction (r_G), as Hull does in the most recent version of his system (1952, Corollaries ii and xv). For if s_G has secondary reinforcing power, then any response that has immediately preceded it will be conditioned to the stimuli present within the preceding short interval. But it is always preceded by the r_G of which it (the s_G) is the physiological result, and by the stimulus—let us call it S_x —which on that occasion evoked the r_G . Therefore, the secondarily reinforcing s_G must always strengthen the $S_x \rightarrow r_G$ bond of which it is the consequence, so that once a stimulus has developed a tendency to evoke an anticipatory goal response, this tendency can never become extinct, i.e. the stimulus (S_x) can never lose its secondary reinforcing power.

IV. LATENT EXTINCTION

In recent years there have been a number of interesting experiments in which an animal is trained to perform a habit, such as running a maze, which brings it to a goal box containing a reward. It is then given a number of experiences of the goal box with no reward in it, and without it having performed the learned responses necessary to reach the goal. Then the reaction potential of the habit is measured, for instance by measuring resistance to extinction. Thus in such an experiment both experimental and control groups of rats would be trained to approach a goal box to receive a reward. After this the experimental rats, but not the controls, would be placed several times directly in the (now unbaited) goal box. Both groups of rats would then be subjected to an extinction procedure. It has been shown in a number of experiments that the experimental group now extinguish more rapidly than the controls, and this procedure has been called 'latent extinction' (Coate, 1956; Deese, 1951; Moltz, 1955; Seward & Levy, 1949).

While such a result is in complete accord with expectancy theories, it causes grave difficulty for $S-R$ theory, in which inhibition is response induced: the reaction potential of any stimulus-response association is increased by reinforcement, reduced by the inhibition produced by response performance. On these assumptions a situation in which reaction potential is reduced without the response having been performed cannot be explained (Gleitman, Nachmias & Neisser, 1954).

Moltz (1957) has recognized this as a major difficulty. However he believes the problem can be met, and $S-R$ theory salvaged, by a very ingenious argument employing the r_G . He distinguishes two types of latent extinction experiment: that described above [his Type I (Deese, 1951; Moltz, 1955; Seward & Levy, 1949)], where the effect of the latent extinction is measured by the reduction in performance in subsequent extinction trials, and a variant (Type II, Moltz & Maddi, 1956), where the effect is shown by using the unbaited goal box to produce secondarily-reinforced learning. In such an experiment the animal is given training on a runway at the end of which it sometimes finds a distinctive goal box (let us suppose it is black) with reward, or a different (white) box with no reward. The experimental rats are then placed directly several times in the black goal box, but with no reward present. The

control group do not have this treatment. In the third part of the experiment the two end-boxes are placed at the ends of opposite arms of a T-maze, and the extent to which animals learn to take the arm leading to the goal box which formerly held the reward is noted. This learning reflects the secondary reinforcing effect which the goal box acquired during the first period of training, when it contained a reward. Moltz & Maddi (1956) found that this learning was less for the experimental group than for the controls, a further instance of latent extinction.

Moltz's (1957) suggested explanation, based on Hull's general behaviour theory, is briefly as follows: In the goal box the full consummatory response, R_G , will occur. Since it is closely associated with reinforcement it will tend to become conditioned to all the stimuli present at the times when it is evoked, i.e. to all the goal box stimuli, so that the goal box stimuli will in the future produce those parts of the R_G which are not incompatible with other ongoing responses—i.e. the fractional anticipatory goal response, r_G . Thus, when the rat is again placed in the goal box, the stimuli there will tend to elicit r_G 's freely and repeatedly. Each fractional goal response will produce a fractional goal stimulus, s_G . This s_G is part of S_G , the full proprioceptive and interoceptive stimulation produced by the consummatory goal response, and since it must therefore have often been associated with reinforcements it has, by virtue of Hull's Corollary ii (1952), secondary reinforcing power.

Therefore at the end of the initial period of training, in each of the two types of experiment, the goal box will have acquired the power to evoke r_G 's freely, each of which then produces an s_G with secondary reinforcing power.

Moltz argued further that in Type I latent extinction some of the stimuli in the goal box will be similar to stimuli in the stem of the maze as the animal approaches the choice-point. Therefore, by the principle of stimulus generalization, these too will evoke r_G 's and their consequent s_G 's. Thus, among the stimuli present when the animal makes its choice will be the s_G , so that it too, like the other stimuli in the stem of the maze, will have the correct response, i.e. turning towards the goal box with the reward, conditioned to it. So the total habit strength operating at the choice-point to make the rat take the arm leading to the goal box will consist of two parts: a link between the $S_{\text{maze-stem}}$ and $R_{\text{goal-box}}$ and a link between s_G and $R_{\text{goal-box}}$. What will happen when the animal is given its period of 'latent extinction' in the goal box? According to Moltz, the cues at the goal will evoke r_G strongly and repeatedly. 'Now if r_G is conceived as a response process that obeys the same behavioral laws as most molar systems, then it would follow that these successive non-reinforced emissions of r_G will produce a sharp reduction in its response strength with respect to cues in the goal location.' This weakening will generalize to the r_G 's evoked in the maze-stem, so that the s_G 's produced there will be weakened or absent. Consequently that part of the habit strength of the correct response which depends on s_G ($s_G \rightarrow R_{\text{goal box}}$) will be weakened or absent, giving the reduced performance found in the Type I latent extinction experiments.

Moltz's argument for Type II latent extinction is similar: at the end of the first period of training the stimuli in the rewarded goal box will freely evoke r_G 's, so that s_G , with its secondary reinforcing capacity, will be almost continuously present. Therefore, when the positive (black) goal box is present at the end of one of the maze arms, the occurrence of s_G when the animal enters it will reinforce the response of

entering the maze arm leading to it, and will therefore build up the habit of turning in that direction. On the other hand: 'Since the reward object is not present... the emission of r_G in the positive goal box following each correct choice will reduce its response strength, as a result of which the intensity of s_G and hence the magnitude of available secondary reinforcement will decrease progressively.' When the r_G ceases to occur the response of going to the formerly positive goal box will no longer be reinforced, and will extinguish. In the groups of rats undergoing latent extinction the decrease in the emission of r_G 's will be partly completed during the latent extinction stage, so that available secondary reinforcement will be less than for the control group when the final maze series begins. Therefore the habit of going to the formerly positive goal box will be more weakly established and sooner extinguished in the experimental group than in the controls, as Moltz & Maddi found (1956).

I have given these arguments fairly fully, so that it should be clear that in each case they rest on the assumption that the ability of goal box stimuli to evoke r_G will decline and disappear when R_G , the full consummatory response, no longer occurs in the goal box. But this is fallacious, because each time the r_G occurs in the presence of goal box stimuli it is followed by a reinforcing state of affairs, in the form of s_G , which has secondary reinforcing power. Therefore each time the sequence $S_{\text{goal box}} \rightarrow r_G$ occurs it is followed by secondary reinforcement and will be strengthened.

It cannot be argued that the reinforcement afforded by s_G is too weak to maintain the $S_{\text{goal box}} \rightarrow r_G$ connexion against the effect of inhibition, since the reinforcing effect of s_G is adequate to build up connexions such as the habit of choosing the maze arm that leads to the formerly positive goal box. Moreover, its reinforcing effect must be greater on the $S_{\text{goal box}} \rightarrow r_G$ connexion, since the delay of reinforcement is so much less between r_G and s_G than between the choice of the correct maze arm and the subsequent occurrence of s_G in the goal box. Nor can it be argued that s_G loses its reinforcing power because it no longer, as Hull puts it, 'occurs repeatedly and consistently in close conjunction with a reinforcing state of affairs' (1952, Corollary ii), since it is part of the consummatory response of eating, so that its secondary reinforcing power should be strengthened every time the animal eats, whatever may happen in the experimental situation.

Thus Moltz's arguments fail because of the errors in the S - R doctrine of secondary reinforcement. He has failed to provide a workable explanation of latent extinction in terms of S - R theory, and the phenomenon remains as evidence favouring an alternative interpretation.

V. REASONING

Deutsch (1956*b*) has thoroughly analysed this topic, which will be referred to only briefly. Maier (1929) found that rats could combine segments of behaviour learned on separate occasions to solve problems, which he regarded as insightful and non-associative. Hull (1935) tried to explain these performances in terms of 'the principles of association between stimuli and responses, particularly as revealed in modern conditioned-reaction experiments'. The sort of situation he analysed is instanced by the following simplified account. A hungry rat is trained to run a segment of maze from a starting box, U , to a goal box with food. On a separate occasion, under a

different drive (thirst), it is also trained to run the two arms of a V for the water which is placed in both goal-boxes. One of the goal-boxes is the starting box, U , of the food segment. When it is subsequently made hungry and placed at the start of the V, it shows insight by choosing the arm leading to U and so on to food. His explanation was that a fractional response, r_u , to the distinctive features of starting box U became conditioned to the drive stimulus of hunger in the first period of training. In the second period, r_u came to be evoked by the start of the V, so that the proprioceptive stimulus s_u would precede entering the arm leading to U and this response would be conditioned to it. On the test, r_u would be evoked both by the start of the V and the drive stimulus of hunger, so making s_u strong, which would favour choice of the arm leading to U . Deutsch points out that the flaw in this and other similar explanations (Hull, 1952) is the failure to consider carefully enough what consequences S - R principles would imply in the different stages of training. When r_u is evoked at the start of the V, this will result in s_u preceding the response of going to the water reward at U , and so it will come to evoke this response. But s_u will just as often precede the choice of the other arm and so will evoke that response just as strongly. Therefore the strengthening of s_u by the drive stimulus of hunger will not favour either response, and cannot account for the rat's successful solution.

VI. DRIVE DISCRIMINATION

If a rat which is both hungry and thirsty is trained on a T-maze with water in one goal box and food in the other, and is subsequently tested with only one of the drives present, it will choose the arm leading to the appropriate reward (Kendler, 1946). As the same pattern of stimuli precedes all choices in training, this is a challenge for S - R theory.

Spence has attempted to meet this challenge by proposing a fractional anticipatory response explanation (Kendler, 1946; Spence 1951 *a, b*; Spence, Bergmann & Lippitt, 1950; Hull, 1952):

During the training series the stimuli in the water box and in the water alley become conditioned to anticipatory drinking acts r_w and, in turn, the proprioceptive components s_w resulting from these anticipatory acts themselves become conditioned to the response of entering and continuing locomotion in this alley. In a similar manner, anticipatory eating responses r_f develop and their cues s_f become conditioned to the response of entering the alley leading to food. During the test series when only one drive is operative, the anticipatory act related to the goal for which the subject is motivated is much stronger than the other and hence produces stronger proprioceptive cues. Thus, when the subject is hungry, proprioceptive cues from anticipatory eating responses are stronger than those for drinking. As these cues become conditioned to the response of entering the food alley, they tend to throw the balance of habit strength to the food-alley response. [Spence, 1951 *a*].

The apparent force of this argument derives from the imprecision of the description 'the response of entering and continuing locomotion in this alley', which confuses two situations. If the cue s_f occurs in the food alley the response of continuing down the alley can be conditioned to it, but not the response of turning towards that alley at the choice point: that has already occurred, and precedes s_f , when the latter is evoked in the alley. When evoked in the alley s_f can only have the effect of supporting the response sequence already chosen.

For the response of turning towards that alley (which alone will allow the stimuli occurring in the alley to become effective in determining behaviour) to be conditioned to s_j , the latter must precede the choice point. But if so, it will precede entering the food alley and entering the water alley equally often. Therefore strengthening s_j will not favour one response against the other (Deutsch, 1956b).

In terms of $S-R$ theory, an animal's responses at any moment are entirely determined by the stimuli and stimulus traces acting on it at that moment. This implies that once the animal is looking down one of the arms its response will be completely determined by the stimuli from that arm, which will act to make it go forward down the arm. Therefore at the stage where it is looking down an arm its response is already determined, and was determined when it chose to look down that arm rather than the other one. And the response made in the maze-stem is just as likely to be the 'wrong' one as the 'right' one.

Ritchie (1959) has attempted to reply to this argument. He states:

This is a shrewd objection, but it must be recognized that it is directed not towards the r_G hypothesis but towards the $S-R$ analysis of choice behaviour. Without an answer to this objection we cannot understand how an animal can perform properly in any discrimination where the animal is required to turn away from the cue which has less *relative* excitatory strength. . . . Thus the question which Deutsch raises asks us to provide an explanation of *VTE* behaviour, and that is not easy. . . . The problem is as important as it is complex, but it is not a problem peculiar to the r_G mediation hypothesis.

Ritchie is correct in recognizing that the $S-R$ analysis of choice behaviour is in question, but wrong if he supposes that an assumption about the occurrence of *VTE* after the animal has turned to one alley or the other would solve the problem. We would still have to explain why the stimuli in the stem, with their confusing effects, did *not* determine the choice, while assuming the occurrence of vicarious trial and error (*VTE*) at the choice point (i.e. the animal would look down both arms in succession and then respond) would produce yet further difficulties. (i) If the animal is looking down one arm then all the stimuli impinging on it will tend to make it proceed further down that arm—yet instead it turns away. This is difficult to reconcile with the assumptions of $S-R$ theory. (ii) Suppose the animal looks first in one direction (say left) and then, because of *VTE*, in the other (right). Then the stimuli from the right arm will be acting on the animal, but only the traces of the stimuli from the left, which will be fading. If two incompatible momentary reaction potentials are simultaneously present, the stronger one will determine the response (Hull, 1952, Corollary xiv), and by the principle of stimulus-intensity dynamism the stimulus actually present, being stronger than the trace, is very much favoured. Therefore there will be a strong tendency for the animal to run down whichever arm it looks down second—if *VTE* prevents it running down the first arm. (iii) Suppose the trace stimulus from the first arm does determine the response. It has always been connected to the response of entering and running down an arm, but the response it must now evoke is that of turning away and going in the opposite direction. (iv) It seems that to make *VTE* 'work', therefore, we will have to assume some sort of central process which measures and compares the two reaction potentials, allowing for fading of the stimulus trace, to decide which response would be made. This is an extra mechanism (over and above $S-R$ connexions) of

some complexity. (v) What brings the *VTE* mechanism into action? The animal is carried along by an *S-R* chain. Why at one point does it fail to respond to the stimuli actually present to it, and instead swing round to look for others? The answer must be: because at this point it can expect stimuli evoking an alternative set of responses possibly leading to its goal. *VTE* occurs only at the choice point, not at any other part of the response sequence. It must therefore reflect the anticipation at this point of an alternative path, or set of stimuli, not immediately present to it. Thus *VTE* would imply the very 'expectancy' considerations the r_G explanation was intended to avoid. (vi). If *VTE* were an essential part of the mechanism of discrimination at a choice point, it must necessarily be a constant feature of the performance, once the discrimination is learnt. But the evidence is that it increases during learning, and then decreases once the task has been learned if it is not too difficult (Goss & Wischner, 1956; Muenzinger, 1938; Tolman, 1938). This was confirmed in an unpublished experiment by the author, similar to one by Seward (1949). A group of eight rats which had been made very familiar with a simple maze with two distinctive goal boxes, were each fed in one of them before a test trial. They all made correct choices on the first test trial ($P = 0.005$) but did not show any *VTE*, or even slowing down or hesitation, at the choice point. If the result were determined by cues in the alleys, *VTE* would have been essential and the result inexplicable.

A second line of attack is offered by the assumption, on which Spence's argument is based, that the hunger drive stimulus would mainly produce r_p , and the thirst stimulus r_w , because 'upon the whole the animal has been accustomed to eat when hungry and drink when thirsty', as Hull puts it (1952). If this were correct and if experience of eating when hungry and drinking when thirsty could be excluded, the animals should then be unable to learn drive discrimination when trained with both drives present, as in Kendler's (1946) experiment. This has been tested by Kendler (1949) who reared rats on an enriched milk diet and then trained and tested them in a double-drive situation similar to his previous experiment. He obtained evidence of significant discrimination. Deutsch (1958), who independently reared rats on mash and performed a similar experiment, also obtained significant discrimination. Thus the rats could discriminate in the absence of an association between the drive stimuli of hunger and thirst and the appropriate fractional responses, which should not have been possible if the r_G explanation were correct.

Ritchie (1959) points out that 'since reinforcement occurs whenever a hungry rat eats, once a chance r_w becomes conditioned to the hunger stimulus it will never be extinguished, but instead will soon reach maximal habit strength'. He considers that this difficulty can be met by making special assumptions.

Kendler (1946) has an alternative explanation for the successful discrimination shown by rats trained with both drives present: a selective principle of the association of drive stimuli and responses. He suggests that 'only those drive stimuli which are themselves reduced become connected to a rewarded response'. This principle is in conflict with the result of an experiment by Spence *et al.* (1950), on learning under irrelevant motivational conditions, where drive stimuli were not reduced. They ran rats, satiated for food and water, in a Y-maze with food on one side and water on the other. After running, the rats were placed with their cage-mates, 'the reinforcement

being provided by the social goal, or possibly... by the sight of the food or water itself (i.e. secondary reinforcement)'. In neither case would the drive stimuli have been reduced, so that, on Kendler's selective principle, no responses could have been attached to them and no learning predicted. Nevertheless on the test trial, when the rats were made hungry or thirsty, they showed a significant number of correct choices.

VII. INCENTIVE MOTIVATION AND LATENT LEARNING

If a rat runs a maze for a reward of bran mash, and sunflower seed is substituted for this accustomed reward, the rats show a disruption of behaviour, dividing 'their time between eating and random searching', and on subsequent days run more slowly and with more errors (Elliott, 1928; Tolman, 1932). If a monkey is allowed to choose a cup in which it has seen banana placed, but finds lettuce instead, it shows 'surprised hunting behavior' and may refuse to eat the lettuce (Tinklepaugh, 1928; Tolman, 1932).

Hull (1952) provides an explanation for the deterioration in the maze-running, though not for the behaviour in the goal box or with the reward, in terms of the incentive motivation factor, K . Can this factor account for latent learning?

K is a function of the weight of food or quantity of other reward. Its operation is defined by postulate VIII (Hull, 1952) as shown in the equation:

$${}_sE_R = D X V_1 X K X {}_sH_R.$$

It is interesting to compare it with drive D , which is also a non-associative factor (Spence, 1951*b*). Hull (1943) states: 'all drives alike are able to sensitize all habits' (p. 247), and the effective drive acting on a given habit is the sum of D and 'other or alien drives active at the time' (pp. 244-5). However, rats could learn to perform different habits according to the nature of the drive active (Hull, 1933; Leeper, 1935) which, as Hull clearly realized, raised the problem of how particular drive states could affect habits differentially. The answer was 'their capacity to call forth... distinctive drive stimuli gives each a considerable measure of distinctiveness and specificity in the determination of action' (p. 241).

K is very similar to D . It is determined by the combined effect of incentive substances (Hull, 1952; Corollary viii), is multiplicative and non-associative. It can be applied to explain the case Hull considers, in which the strength of a single dominant habit varies with change in the reward. But what happens in the more complex and interesting instances of latent learning where there is no dominant habit and the incentive must determine a choice? Seward (1949) allowed rats to explore a T-maze without food, then fed them in one goal-box. On the next trial 28 of 32 rats went directly to that goal-box. Kimball, Kimball & Weaver (1953) obtained similar results using a multiple T-maze and taking precautions to prevent the run to the goal-box being rewarded.

K cannot account for such results. Its multiplicative effect would have strengthened all habits present in the situation, so that the directive effect of the incentive remains unexplained.

Spence has suggested that K is determined by the r_G : 'through the intensity of its trace, which presumably increases with the increased strength of r_G as the latter becomes more strongly conditioned, s_G determines the nonassociative factor, K ,

(Spence, 1951*b*, p. 273). This carries us no further; K is still nonassociative, while in an experiment such as that of Kimball *et al.* there has been no opportunity for the conditioning of r_G to occur.

VIII. CONCLUSION

This examination of the more plausible S - R solutions for the more important problems shows that as yet, despite the concentrated ingenuity of numerous exponents, no tenable explanations of the more important problems have been produced.

With considerable justification the early behaviourists rejected explanations in terms of 'expectancy' and 'purpose', since such ideas were used in a grossly anthropomorphic and subjective way, and their emphasis on objective observation and verification has been very valuable. Today, however, with modern knowledge of cybernetics, servomechanisms and computers, we know that devices to perform quasirational tasks can be constructed in a variety of ways, and it is possible to define the requirements of an expectancy theory in a completely objective way. These requirements would be as follows. (i) The organism can store information about the outcome of a response, i.e. the stimulus situation eventuating from a given response in a given situation. (ii) When the animal is seeking a goal this information can be employed to determine the immediately effective sub-goal (e.g. which of two paths will be sought at a choice-point). (iii) The stored information can be altered in ways which do not require performance of a response.

'Storage of information' is accomplished by any central nervous change in the mechanisms controlling behaviour which will alter future behaviour in a relevant way. A number of attempts have been made to describe mechanisms which might meet these requirements (Deutsch, 1956*a*; Hebb, 1949; Milner, 1957; Olds, 1954).

This work was done while in receipt of a grant from the Medical Research Council. I would like to thank Prof. R. C. Oldfield for facilities afforded, and Dr J. A. Deutsch for many helpful discussions.

REFERENCES

- COATE, W. B. (1956). Weakening of conditioned bar-pressing by prior extinction of its subsequent discriminated operant. *J. comp. Physiol. Psychol.* **49**, 135-8.
- DEESE, J. (1951). The extinction of discrimination without performance of the choice response. *J. Comp. physiol. Psychol.* **44**, 362-6.
- DEUTSCH, J. A. (1956*a*). A theory of insight, reasoning and latent learning. *Brit. J. Psychol.* **47**, 115-25.
- DEUTSCH, J. A. (1956*b*). The inadequacy of the Hullian derivations of reasoning and latent learning. *Psychol. Rev.* **63**, 389-99.
- DEUTSCH, J. A. (1958). Double drive learning in rats without previous selective reinforcement. *Quart. J. Exp. Psychol.* **10**, 207-10.
- ELLIOTT, M. H. (1928). The effect of change of reward on the maze performance of rats. *Univ. Calif. Publ. Psychol.* **4**, 19-30.
- GLEITMAN, H., NACHMIAS, J. & NEISSER, U. (1954). The S - R reinforcement theory of extinction. *Psychol. Rev.* **61**, 23-33.
- GOSS, A. E. & WISCHNER, G. J. (1956). Vicarious trial and error and related behaviour. *Psychol. Bull.* **53**, 35-54.
- HEBB, D. O. (1949). *The Organization of Behavior*. New York: Wiley.
- HILGARD, E. R. & MARQUIS, D. G. (1940). *Conditioning and Learning*. New York: Appleton-Century-Crofts.

- HULL, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.* **38**, 487-506.
- HULL, C. L. (1933). Differential habituation to internal stimuli in the albino rat. *J. Comp. Psychol.* **16**, 255-73.
- HULL, C. L. (1935). The mechanism of the assembly of behavior segments in novel combinations suitable for problem solution. *Psychol. Rev.* **42**, 219-45.
- HULL, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century Co.
- HULL, C. L. (1952). *A Behavior System*. Yale University Press.
- KENDLER, H. H. (1946). The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. Exp. Psychol.* **36**, 212-20.
- KENDLER, H. H. (1949). An experimental examination of the non-selective principle of association of drive-stimuli. *Amer. J. Psychol.* **62**, 382-91.
- KIMBALL, R. C., KIMBALL, L. T. & WEAVER, H. E. (1953). Latent learning as a function of the number of differential cues. *J. Comp. physiol. Psychol.* **46**, 274-80.
- LEEPER, R. (1935). The role of motivation in learning: a study of the phenomenon of differential motivational control of the utilization of habits. *J. Genet. Psychol.* **46**, 3-40.
- MAIER, N. R. F. (1929). Reasoning in white rats. *Comp. Psychol. Monog.* **6**, no. 29.
- MILLER, N. E. (1935). A reply to 'Sign-Gestalt or conditioned reflex?' *Psychol. Rev.* **42**, 280-92.
- MILNER, P. M. (1957). The cell assembly: Mark II. *Psychol. Rev.* **64**, 242-52.
- MOLTZ, H. (1955). Latent extinction and the reduction of secondary reward value. *J. Exp. Psychol.* **49**, 395-400.
- MOLTZ, H. (1957). Latent extinction and the fractional anticipatory response mechanism. *Psychol. Rev.* **64**, 229-41.
- MOLTZ, H. & MADDI, S. R. (1956). Reduction of secondary reward value as a function of drive strength during latent extinction. *J. Exp. Psychol.* **52**, 71-6.
- MUENZINGER, K. F. (1938). Vicarious trial and error at a point of choice: I. A general survey of its relation to learning efficiency. *J. Genet. Psychol.* **53**, 75-86.
- OLDS, J. (1954). A neural model for sign-Gestalt theory. *Psychol. Rev.* **64**, 59-72.
- PAVLOV, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.
- RITCHIE, B. F. (1959). Explanatory powers of the fractional antedating response mechanism. *Brit. J. Psychol.* **50**, 1-15.
- SEWARD, J. P. (1949). An experimental analysis of latent learning. *J. Exp. Psychol.* **39**, 177-86.
- SEWARD, J. P. & LEVY, N. (1949). Sign learning as a factor in extinction. *J. exp. Psychol.* **39**, 660-8.
- SHERRINGTON, C. S. (1906). *The Integrative Action of the Nervous System*. London: Constable.
- SPENCE, K. W. (1951a). Theoretical interpretations of learning. Chapter 18 in *Handbook of Experimental Psychology*. Ed. by S. S. Stevens. New York: Wiley.
- SPENCE, K. W. (1951b). Theoretical interpretations of learning. Chapter 8 in *Comparative Psychology*. Ed. by C. P. Stone. New York: Prentice-Hall.
- SPENCE, K. W. (1956). *Behavior Theory and Conditioning*. Yale University Press.
- SPENCE, K. W., BERGMANN, G. & LIPPITT, R. (1950). A study of simple learning under irrelevant motivational-reward conditions. *J. Exp. Psychol.* **40**, 539-51.
- TINBERGEN, N. (1951). *The Study of Instinct*. Oxford: The Clarendon Press.
- TINKLEPAUGH, O. L. (1928). An experimental study of representative factors in monkeys. *J. Comp. Psychol.* **8**, 197-236.
- TOLMAN, E. C. (1932). *Purposive Behavior in Animals and Men*. New York: Appleton-Century.
- TOLMAN, E. C. (1938). The determiners of behavior at a choice point. *Psychol. Rev.* **45**, 1-41.
- WATSON, J. B. (1924). *Psychology from the standpoint of a Behaviorist*. Philadelphia and London: Lippincott. Second edition.
- ZENER, K. (1937). The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response. *Amer. J. Psychol.* **50**, 384-403.

(MS. received 20 February 1959)

CONSTRUCTION OF A PSYCHOLOGICAL SCALE OF WEIGHT FROM FRACTIONATION DATA

BY R. G. CHATTERJEA AND RAMANATH KUNDU

Psychology Laboratory, Calcutta University, Calcutta, India

A subjective scale of weight, by the method of fractionation, has been attempted with graphical and analytical procedures. The comparison weights were presented in two ways—in the first they accompanied the standard weight serially and in the second each standard weight was accompanied by all the comparison weights. A weight of 63 g. was found to be judged half as heavy as a physical weight of 100 g. The subjective weight corresponding to the physical weight of 100 g. was taken as the unit of subjective weight and it was termed as 1 veg. In terms of this unit a veg scale of subjective weight has been constructed. An empirical equation relating to the physical weight in grammes and the subjective weight in vegs has also been deduced.

I. INTRODUCTION

The fractionation method has been widely used for the construction of scales and has been extended to different sense modalities. The veg scale for lifted weights has been worked upon by various methods. The scales obtained by the fractionation method are those of Guilford & Dingman (1954), and Warren & Warren (1956). The construction of a psychological magnitude (subjective weight) by the same method has been treated by Harper & Stevens (1948). Earlier studies relating to the fractionation method under different sense modalities have been described by Stevens & Galanter (1957). The present paper deals with an experiment using the same method as that of Harper & Stevens (1948) with a slight variation in the procedure. An attempt has also been made to show whether this variation in procedure has any bearing on the results.

II. PLAN OF INVESTIGATION

(a) *Subjects and materials*

The subjects in the present experiments were twenty-five students of the post-graduate and undergraduate classes in Psychology. The weights were contained in pharmacists' pill boxes of equal shape and size. Five series of weights were tried. Each series comprised of a standard and four comparison weights (Table 1).

Table 1. *Standard and comparison weights*

Standard weight (g.)	Comparison weight (g.)
40	24, 28, 32, 36
80	48, 56, 64, 72
120	72, 84, 96, 108
160	96, 112, 128, 144
200	120, 140, 160, 180

(b) *Experimental design*

The procedure consisted of two parts. (i) Presenting the comparison weights serially corresponding to each standard weight, and (ii) presenting all the comparison weights with each standard weight. In the former case, for example, a standard weight of 40 g. was accompanied by comparison weights of 24, 28, 32 and 36 g.—and the sequence continued with other series.

But in the latter case a standard weight of 40 g. was presented with all the comparison weights from 24 to 180 g. as listed in Table 1. When judgement with 40 g. was obtained, the next standard weight was presented with the other comparison weights, leaving aside the one selected by the subject and the sequence continued with other standard weights.

In (i) the subjects were required to select the comparison weight that 'feels half as heavy as the standard'. Here the operation was to select any one from the four comparison weights. But in (ii) the subjects lifted the standard first and then tried one by one all the comparison weights to find out one that 'feels half as heavy as the standard'. In this case, the operation was rather complex in that the subject had to lift the standard and then any one of the comparison weights till he could find the weight that felt half as heavy as the standard. This procedure apparently required much time and needed too many repetitions.

The working table was sufficiently illuminated. The comparison weights were numbered on the bottom surface of the pill boxes. Since the pill boxes were of equal shape and size, all possible clues were controlled. The presentation of the series with procedure (i) is termed as 'serial presentation', while that with (ii) is called 'haphazard presentation'.

III. RESULTS AND DISCUSSION

The medians of the twenty-five judgements made for each of the series and their corresponding logarithmic transformations are shown in Table 2.

Table 2. *Medians for two series of twenty-five judgements each*

Standard weight	log X	Weight judged half as heavy as the standard			
		Serial presentation (median) (Y)	log Y	Haphazard presentation (median) (Z)	log Z
40	1.602	27.8	1.444	27.8	1.444
80	1.903	55.7	1.746	48.3	1.683
120	2.079	83.8	1.923	83.8	1.923
160	2.204	111.6	2.048	96.4	1.984
200	2.301	120.5	2.081	120.0	2.079

The median judgements in serial presentation (Y) practically coincide with those of haphazard presentation (Z) except in the case of two standard weights—80 and 160 g. But the magnitude of variation between the median judgements of the serial presentation and the haphazard presentation with respect to these standard weights are not sufficiently high. So we may infer that variation in the presentation (serial and haphazard) of the weights does not produce any variation in the judgements of the subjects. This result is of considerable importance. From the absence of any difference in results with these methods of presenting the weights, it is obvious that our proposed psychological scale of weight from two series of fractionation data will not differ significantly. Hence separate treatment of the data of series Y and Z seems to be unnecessary and only the judgements of haphazard presentation will henceforth be considered.

A scale of subjective weight from fractionation data for lifted weights may be constructed either graphically or analytically. We propose to consider the construction of such a scale by the graphic method first.

In order to determine a relationship between physical weight and subjective weight, the log values of Table 2 have been plotted in Fig. 1. This curve gives the

logarithm of the weight judged subjectively to be half as heavy, plotted against the logarithm of the corresponding physical weight. The weight judged half as heavy as a physical weight of 100 g. is found to be (from Fig. 1) 63 g. Now if we define, in the manner of Harper & Stevens (1948), the unit of subjective weight as the weight experienced when one lifts 100 g. and call it 1 veg, we may say that physical weight of 100 g. corresponds to subjective weight of 1 veg. Since a weight of 63 g. is found to be judged half as heavy as that of 100 g., we may assign 0.5 veg (subjective weight) to the corresponding physical weight of 63 g. We can similarly find physical weights corresponding to the subjective weights of 2 vegs, 4 vegs, etc. Table 3 shows certain physical weights in grammes and corresponding subjective weights in veg.

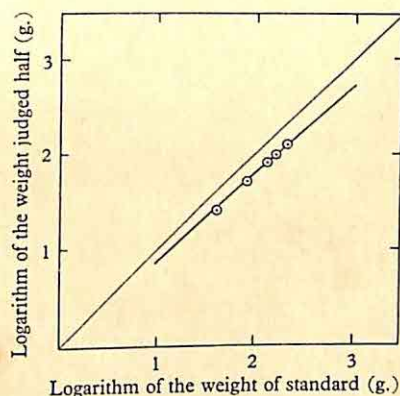


Fig. 1. Plot of fractionation data.

Table 3. *Physical weight and subjective weight*

Physical weight (g.) (<i>W</i>)	$\log W$	Subjective weight in vegs (<i>V</i>)	$\log V$
63	1.8	0.50	1.7
100	2.0	1.00	0
158.5	2.2	2.00	0.3
251.2	2.4	4.00	0.6
398.1	2.6	8.00	0.9
631.0	2.8	16.00	1.2
1000.0	3.0	32.00	1.5

The data of Table 3 can be graphically presented to represent a veg scale of subjective weight.

Now to construct a scale of subjective weight analytically, an empirical equation relating to the physical weight in gramme W and the number of vegs V can be deduced. In order to deduce such an equation, a graphical plot of $\log V$ against $\log W$ (Table 3) was made as shown in Fig. 2. The figure shows that the points come very close to a straight line. Hence the procedure is much simplified.

The relation between V and W can now be expressed as

$$\log V = m \log W + C,$$

where m and C are to be determined. Now m can be easily obtained from the slope of the straight line of Fig. 2, as

$$\begin{aligned} m &= \frac{\log V_2 - \log V_1}{\log W_2 - \log W_1} \\ &= \frac{0.75 - 0}{2.5 - 2} = \frac{0.75}{0.5} = 1.5. \end{aligned}$$

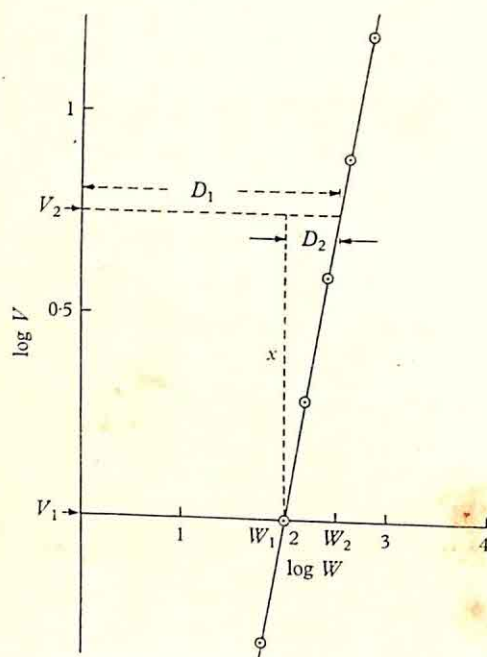


Fig. 2. The logarithmic plot of the veg scale of subjective weight.

V_2 and V_1 are the subjective weights corresponding to the physical weights W_2 and W_1 respectively. V_2 and V_1 and W_2 and W_1 are indicated in Fig. 2. C was obtained from the intersection made by the straight line on the $\log V$ axis and was found to be -3 . Hence the equation becomes

$$\log V = 1.5 \log W - 3.$$

With the help of this equation one can easily get the number of vegs V corresponding to any physical weight W .

The above equation, which was derived in a very simple and straightforward manner, will be, it is hoped, more useful in the practical field. We have also deduced an equation following Harper and Stevens's (1948) method and this is given by

$$\log V = 13.5 \log (1.88 + \log W) - 7.96.$$

IV. CONCLUSIONS

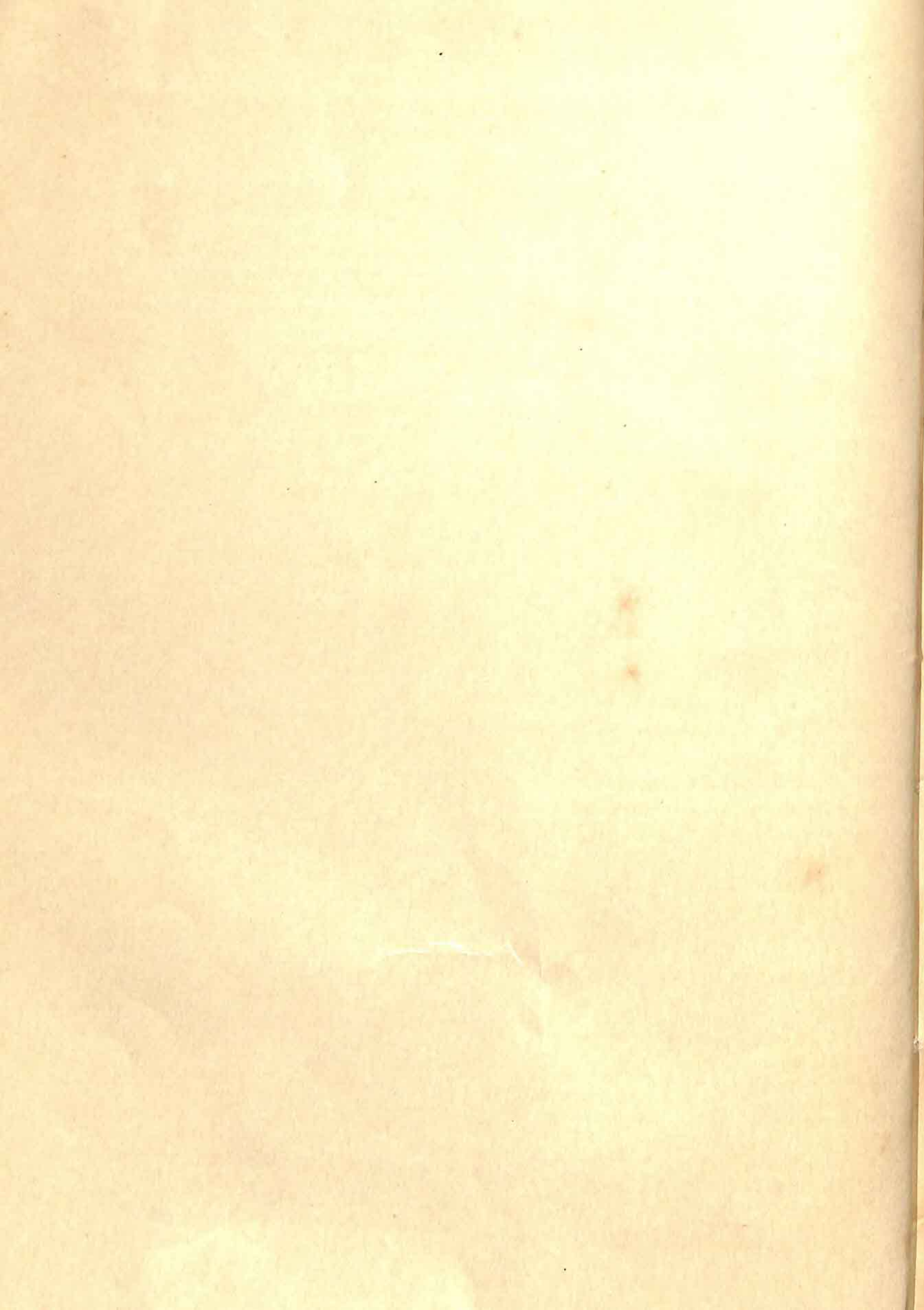
The psychological scale of weight determined by the fractionation method slightly differs from the one shown by Harper & Stevens (1948). Though we presented the comparison weights in two ways, the estimations of the subjects, as shown by the median judgements, were not significantly different. The variation in the presentation

of the comparison weights, therefore, does not bring any significant variation in the judgements. The complex equation of Harper & Stevens has been presented in a more simple form.

REFERENCES

- GUILFORD, J. P. & DINGMAN, H. F. (1954). A validation study of ratio judgements methods. *Amer. J. Psychol.* **67**, 385-410.
- HARPER, R. S. & STEVENS, S. S. (1948). A psychological scale of weight and a formula for its derivation. *Amer. J. Psychol.* **61**, 343-51.
- STEVENS, S. S. & GALANTER, E. H. (1957). Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.* **54**, 377-411.
- WARREN, R. M. & WARREN, R. P. (1956). Effect of the relative volume of standard and comparison object on half-heaviness judgements. *Amer. J. Psychol.* **69**, 640-3.

(Manuscript received 5 January 1959)



DEGREE CLASS AND ATTAINMENT IN SCIENTIFIC RESEARCH

By LIAM HUDSON

The Psychological Laboratory, University of Cambridge

The most important criterion for the award of a grant for post-graduate research in science appears at present to be the class obtained by the candidate in his final B.A. or B.Sc. degree. As a preliminary check on the validity of this criterion, the degree classes gained by four groups of eminent British scientists have been ascertained. These groups comprise Fellows of the Royal Society and Doctors of Science who have graduated from either Oxford or Cambridge during the period 1920-39. The degree classes obtained by members of comparable groups of research scientists who have not as yet been elected to Fellowship of the Royal Society or awarded the Doctorate of Science are used as a basis of comparison.

The results of these comparisons indicate: (i) that neither the Cambridge F.R.S.'s nor the Cambridge D.Sc.'s had better degree classes than those of their respective control groups; (ii) that although Cambridge F.R.S.'s were roughly three times as numerous as those from Oxford, they were far less likely to have first class degrees and (iii) that the F.R.S.'s of both Universities gained considerably better degrees than the D.Sc.'s.

Some implications of these results are discussed.

I. INTRODUCTION

The selection of graduates for training in scientific research is a matter of very real contemporary importance. Yet no systematic attempt appears to have been made, in this country at least, either to formulate the criteria governing the award of grants for post-graduate training, or to assess their validity by follow-up studies. Nevertheless, it is widely known that great importance is at present attached to the *class of degree* gained by a candidate for a research award. Although this is seldom, if ever, the sole criterion employed, it is typically an over-riding consideration. For instance, it is very unusual for a candidate with a good second-class degree to be preferred to one with a first-class degree, or, in cases in which the second class is divided, for a candidate placed in the lower division to be preferred to one placed in the upper division. Further, it is extremely rare for a candidate who was placed in the third class to be awarded a research grant under any circumstances whatsoever.

The basic problem in assessing the present selection method is, therefore, to ascertain whether a positive relation exists between degree class and ability in scientific research.

II. METHOD

Although it would be most desirable to undertake follow-up studies of research workers who have been awarded grants on the basis of contemporary selection procedures, it is possible that some valuable information might be obtained by means of a retrospective inquiry. That is to say, an attempt might be made to ascertain the degree classes of those who have subsequently gained eminence in scientific research. It is true that such an inquiry presents certain difficulties. First, there is no practicable way of assessing those individuals who might have succeeded in scientific research, had they in fact had the opportunity to undertake it. Secondly, there is no generally accepted criterion of 'success' in scientific work. Although it is hardly possible to meet the first objection, the second may be overcome, to a very considerable extent, by employing such criteria as election to the Royal Society or award of the degree of Doctor of Science. In both cases, a considerable body of research, some of it at least of outstanding quality, is an indispensable qualification. The existence of these two criterion groups makes a preliminary check on the relation between degree class and research ability a practical possibility.

The method employed was thus a comparison of the degree classes of research scientists who subsequently became Fellows of the Royal Society or Doctors of Science with the degree classes of those who did not.

If selection for research is predominantly in terms of degree class, and if degree class is positively related to research ability, we may predict that the proportion of good degrees will be significantly higher among Fellows of the Royal Society and Doctors of Science than among scientists who have not achieved either distinction. If, on the other hand, degree class bears no relation to research ability, no such discrepancy should be in evidence. (It is of course not necessary to assume that every F.R.S. and every D.Sc. has greater research ability than every non-F.R.S. and every non-D.Sc., but simply that the average ability among F.R.S.'s and D.Sc.'s is higher than the average ability among non-F.R.S.'s and non-D.Sc.'s.)

Throughout the present study, attention was restricted to the Oxford Final Honours School and the Cambridge Tripos, and all the information used was obtained from published records. The choice of criterion groups was restricted to include only those living F.R.S.'s and D.Sc.'s who took their first degree from either Oxford or Cambridge during the period 1920-39. It was assumed that standards both of examination and of acceptance for post-graduate research remained reasonably uniform during this period, whereas some fluctuation may well have occurred during the wars. It was further thought that the inclusion of those who graduated before 1914 might well raise additional issues of no relevance to modern selection problems.

The results from Oxford and Cambridge respectively have been treated separately throughout. We thus have four distinct criterion groups:

- (i) Fellows of the Royal Society who graduated at Oxford;
- (ii) Fellows of the Royal Society who graduated at Cambridge;
- (iii) Doctors of Science in the University of Oxford;
- (iv) Doctors of Science in the University of Cambridge.

These four groups are completely independent in that no scientist belongs to more than one group; if a scientist happened to be both an F.R.S. and a D.Sc., he was always classified as an F.R.S.

In gathering balanced samples for the first two control groups, each F.R.S. was matched as precisely as possible with a research scientist of the same sex, who read the same subject, who graduated in the same year, but who has not yet been elected to the Royal Society. A similar procedure was used with the D.Sc.'s. In order that the groups should be of sufficient size to justify tests of statistical significance, the sampling procedure was repeated twice in the case of Oxford F.R.S.'s and Cambridge D.Sc.'s, and three times in the case of Oxford D.Sc.'s. The Cambridge non-F.R.S. control group is therefore the same size as the group of Cambridge F.R.S.'s; the Cambridge non-D.Sc. control group is twice as large as the group of Cambridge D.Sc.'s; the Oxford non-F.R.S. control group is twice as large as the group of Oxford F.R.S.'s; and the Oxford non-D.Sc. control group is three times as large as the group of Oxford D.Sc.'s. Apart from these requirements the sampling was entirely random.

III. RESULTS

The degree classes gained by the members of the four criterion groups and their respective control groups are shown in Tables 1-4. For the Cambridge groups (Tables 1 and 2), classes are given for both Parts of the Tripos if both were taken; for the Oxford groups (Tables 3 and 4) the class given is always that gained in the Final Honours School.

The following conclusions may be drawn from these results. First, neither the Cambridge F.R.S.'s nor the Cambridge D.Sc.'s have better degree classes than the members of their respective control groups. (Indeed the degree classes of the Cambridge D.Sc.'s appear if anything to be somewhat worse.) The s.e. of the difference between the proportions being 0.068 in the case of the F.R.S.'s and 0.090 in the case of the D.Sc.'s, it is most unlikely that a significant difference in the proportions has been disguised by inadequate sampling. Secondly, although Oxford

F.R.S.'s are far less numerous than those from Cambridge (in the proportion of 31:91), they are significantly more likely to have gained a first-class degree ($P = 0.005$). Thirdly, the degree classes gained by the F.R.S.'s of both Universities are significantly better than those gained by D.Sc.'s ($P = 0.01$).

Table 1. *Tripes Classes obtained by Fellows of the Royal Society and their controls who graduated from Cambridge*

Class		F.R.S.		Non-F.R.S.	
Part I	Part II	<i>n</i>	%	<i>n</i>	%
1	1	50	55	60	66
2 or 3	1	8	9	6	7
1	2 or 3	10	11	8	9
2 or 3	2 or 3	7	8	8	9
1	—	12	13	6	7
2 or 3	—	4	4	3	3
All		91		91	

Table 2. *Tripes Classes obtained by Doctors of Science of the University of Cambridge and their controls*

Class		D.Sc.		Non-D.Sc.	
Part I	Part II	<i>n</i>	%	<i>n</i>	%
1	1	19	42	53	59
2 or 3	1	4	9	8	9
1	2 or 3	7	16	12	13
2 or 3	2 or 3	5	11	14	16
1	—	3	7	1	1
2 or 3	—	7	16	2	2
All		45		90	

Table 3. *Classes in the Final Honours School obtained by Fellows of the Royal Society and their controls who graduated from Oxford*

Class	F.R.S.		Non-F.R.S.	
	<i>n</i>	%	<i>n</i>	%
1	30	97	36	58
2	1	3	22	35
3	—	—	4	6
4	—	—	—	—
All	31		62	

Table 4. *Classes in the Final Honours School obtained by Doctors of Science of the University of Oxford and their controls*

Class	D.Sc.		Non-D.Sc.	
	<i>n</i>	%	<i>n</i>	%
1	9	69	24	62
2	1	8	13	33
3	1	8	2	5
4	2	15	—	—
All	13		39	

IV. DISCUSSION

The first point of interest is the fact that scientists from Cambridge who have been elected to the Royal Society (or who have been awarded the degree of D.Sc.) do not in general possess better undergraduate academic records than scientists of comparable background who have so far failed to attain either distinction.

If our result is accepted as approximately accurate, it would appear that class in the Tripos bears no direct relation to subsequent eminence in scientific research as assessed by either of our two criteria. Indeed the evidence suggests that the reason why there are relatively few Cambridge F.R.S.'s and D.Sc.'s with seconds and thirds is not that graduates placed in these classes necessarily lack research ability, but rather that they are less frequently given the opportunity to display it.

This interpretation could be restated in the form of the following pair of inferences. (1) During the period 1920-39, a science research student at Cambridge who had been placed in the second class had the same chance of eventually being elected to the Royal Society (or being awarded a D.Sc.) as a similar research student who had been placed in the first class. That is to say, when once accepted for research by whatever criteria, degree class is irrelevant to ultimate achievement in research. (2) During the same period, degree class was irrelevant to research ability whether the student was in fact accepted as a research student or not.

It will be seen that Inference 1 is unobjectionable but only tangentially related to the original problem, while Inference 2 is both directly relevant and open to at least two kinds of justifiable criticism.

These criticisms are as follows. First, the evidence does not exclude the possibility of a positive relation between degree class and research ability in graduates placed in the lower division of the second or in the third class. That is to say, although degree class may have no predictive value in the case of research candidates who have gained firsts and upper seconds, it may still have predictive value in the case of candidates who have lower seconds and thirds. It will be borne in mind that our inquiry provides no check on graduates who might wish to have undertaken research, but who were for any reason not accepted to do so. Secondly, it is possible in the period under consideration, that valid criteria other than degree class were used in the selection of research candidates with seconds and thirds, which were not used in the selection of candidates with firsts. Such might include recommendations from tutors, supervisors and heads of departments. Casual assessments of this kind tend both to be highly unreliable and considerably influenced by the level of the candidate's academic achievement. Their predictive value would not appear therefore to have been very great. None the less, their influence cannot now be measured, so Inference 2, even if restricted to graduates with firsts and seconds, must be viewed with considerable caution.

The second interesting feature of the results is the difference revealed between the classes obtained by Fellows of the Royal Society having Oxford and Cambridge degrees respectively. As we have seen, F.R.S.'s who graduated from Oxford are significantly more likely to have obtained first-class degrees. If we accept Fellowship of the Royal Society as an adequate criterion of research ability, this discrepancy is open to at least three rival explanations, not necessarily mutually incompatible.

(1) That the Oxford Final Honours School is, for whatever reasons, a more valid index of potential research ability than the Cambridge Tripos. (2) That the standards of the Final Honours School at Oxford are lower than those of the Cambridge Tripos. (3) That, for reasons unrelated to research ability, election to the Royal Society is more likely in the case of a Cambridge than an Oxford graduate.

Fortunately, it is possible to check the validity of (2) and (3). For if either explanation is true (or if both are true together), we may predict that simply in terms of total numbers, the proportion of F.R.S.'s from Oxford to those from Cambridge would be smaller than the proportion of Oxford science graduates to Cambridge science graduates. If, on the other hand, both explanations are false, the proportions should be approximately the same. These proportions were accordingly calculated.

The proportion of Oxford F.R.S.'s to Cambridge F.R.S.'s is approximately 1:2.9. The proportion of scientists (mathematicians, natural scientists and engineers) who graduated from Oxford between 1920-39 as compared with those who graduated from Cambridge during the same period is approximately 1:2.7. However, it might still be argued that it is not the total number of science graduates, but rather the number given the opportunity for research that is the crucial consideration. Consequently, the sizes of the populations engaged in scientific research at Oxford and Cambridge during the 1920-39 period were ascertained and compared. The resulting proportion, roughly 1:2.8, is almost the same. It may therefore be said that explanations (2) and (3) are not supported by the available evidence, and that explanation (1) alone remains.

Although no direct evidence bearing on the predictive superiority of the Oxford Final Honours School is available, certain well-known differences between the Oxford and Cambridge systems of examination may be adduced. In the first place, candidates in certain of the Honours Schools at Oxford devote two years study exclusively to the subject in which they are examined, and in which, presumably, they will wish to undertake research. In Cambridge, on the other hand, Part II of the Tripos is usually taken in a single year. In the second place, the Oxford Final Honours School normally includes a *viva voce* examination, which may be of decisive importance in marginal cases. In the third place, it is sometimes claimed that the Oxford marking system is less rigid, allowing greater scope to evidence of originality at the expense of precise knowledge. It is also perhaps the case that the reputation of a first-class degree is somewhat higher at Oxford than at Cambridge, and that particular care is therefore devoted to maintaining the standards of examination.

It is perhaps worth noting that this finding is in apparent contradiction to the current trend in favour of rigidly structured, factual examinations, and seems to favour a more flexible, holistic approach. The validity of such a system obviously hinges on the perspicacity of the examiners, and this may well explain both its success in the present context and its relative failure elsewhere.

The third outcome of this study is the difference revealed between the degree classes achieved by Fellows of the Royal Society and Doctors of Science respectively. This finding might perhaps be taken as an indication of a difference in the standards of research ability necessary for success at the two levels. On the other hand, the discrepancy could equally well be explained as reflecting a difference in the means by which a scientist becomes a Doctor of Science as opposed to a Fellow of the Royal

Society. The degree of Doctor of Science is obtained as a result of a personal application on the part of the candidate, whereas Fellowship of the Royal Society is not. Thus it is quite possible that a large proportion of scientists who are D.Sc.'s but not F.R.S.'s applied for the D.Sc. degree precisely because their undergraduate academic record seemed poor as compared with their achievement in research.

These findings, and the doubt they cast on the validity of the present criteria of selection for research, serve to emphasize the need for a clear-cut distinction between the educational and the selective functions of University examinations. The Oxford and Cambridge B.A. examinations have presumably been evolved as tests of competent knowledge in the subject which has been studied, including, in the case of scientific subjects, practical competence in the conduct of experiments. Although 'competence' in this context may be assumed to include some evidence of originality, such examinations were not evolved as tests of a student's capacity for original research. The assumption that one examination can serve both educational and selective functions cannot simply be justified *a priori*. The present practice demands independent validation; and the preliminary findings of this study, for the Cambridge system at least, are thoroughly discouraging.

It would perhaps be possible to collect more information of relevance to the present case. A detailed study could be made of the academic histories of individual F.R.S.'s and D.Sc.'s; and some clearer idea could be gained of methods used, during the period 1920-39, in selecting students for scientific research. However, this would not only be a lengthy undertaking, involving large quantities of anecdotal evidence, but it would also be critically dependent on the co-operation of the Royal Society.

In fact it is difficult to see how a proper assessment of rival selection techniques can be reached in the absence of an extensive and inconvenient social experiment. As a means of accumulating the necessary evidence, it might perhaps be possible to institute trial periods of research, lasting six months or a year. Tutors' assessments of the candidate could then be based on his actual performance in research; and, as long as neither criterion was allowed to influence selection for the trial period, the predictive value of degree class and tutors' recommendations could be directly compared.

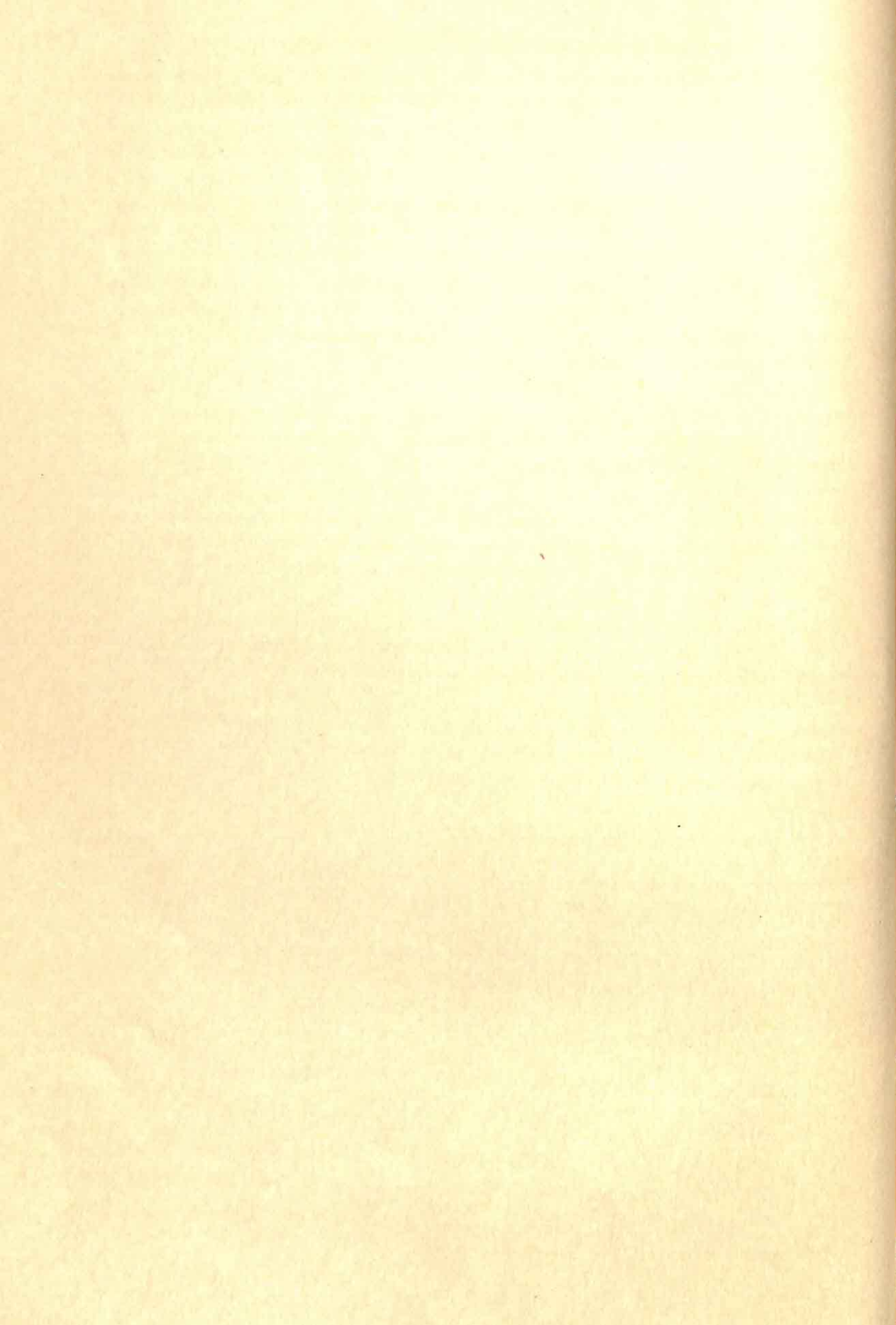
In the meanwhile, there seem to be adequate grounds for two practical suggestions. First, that faith in Oxford and Cambridge degree class as a predictive index of potential research ability in science should be viewed with scepticism. Secondly, that if examinations of academic competence are to be used, for lack of any alternative, for the selection of research scientists, then the Oxford system should be expected to reject fewer candidates of potentially outstanding ability than the Cambridge one.

I am extremely grateful to Prof. O. L. Zangwill for his detailed help at every stage of this inquiry, to Prof. F. G. Young and Dr A. W. Heim for their encouragement and criticism, to Mr Charles Gross for his share in the first part of the work, and to the Editors of *Nature* for permitting me to reproduce material which first appeared in that journal.

REFERENCES

- GROSS, C. & HUDSON, L. (1958). Undergraduate academic record of Fellows of the Royal Society. *Nature, Lond.*, **182**, 787.
- HUDSON, L. (1958). Undergraduate academic record of Fellows of the Royal Society. *Nature, Lond.*, **182**, 1326.

(Manuscript received 21 January 1959)



INHELDER AND PIAGET'S *THE GROWTH OF LOGICAL THINKING*†

II. A LOGICIAN'S VIEWPOINT‡

By CHARLES PARSONS

Society of Fellows, Harvard University

The purpose of this discussion is to examine the formal models which Piaget uses to describe the operational stages. At the concrete level, the subject is said to be able to classify and order and to combine and compare classes and relations according to certain systems. These systems are the eight groupings (*groupements*) set forth in chapters II and III of the *Traité de Logique* (1949), referred to below as TL. In *The Growth of Logical Thinking* (GLT), the authors are not concerned to show the operation of these groupings in detail. I should say that the only sense in which it is shown that they are a *complete* description of concrete logic is that a number of further operations and principles are shown to appear only at the formal level, and it is said that they are best characterized by reference to the 'integrated structure' (*structure d'ensemble*) of propositional logic. The result is to leave open a certain question. The groupings of relations in particular cover only those which satisfy certain special conditions (equivalence relations or ordering relations) and others directly logically related to them. Such relations derive interest from the fact that they lead most readily to somewhat ramified structures and from their importance in applications. But relations which do not fit into these structures occur constantly in ordinary life and in science, and they are susceptible of some logical operations such as composition, conversion, and negation.⁽¹⁾ [Notes appear at the end of this review.] Is Piaget simply picking out the types of relations which yield the richest algebraic structures, or which play the most important role in the cognitive acquisitions of the stage, or does he claim a deeper completeness for the system of eight groupings? In the latter case, the claim is obscure and doubtful.

From the age of 11 onward, subjects are said to develop formal operations. According to Piaget, what this means is that the subjects use the logical forms of the propositional calculus and understand (in concrete cases) the logical relations formalized in the propositional calculus. They operate with *propositions* and therefore can see what actually occurs as one of a number of possibilities and deduce and verify the consequences of hypotheses. They develop a number of 'operational schemata' by means of structures involved in the system of propositional operations. Examples are proportions, mechanical equilibrium, combinatorial operations (pp. 307-29). [All references are to *The Growth of Logical Thinking* unless otherwise stated.]

Although Piaget uses the standard notation of the propositional calculus, it turns out that he vacillates somewhat in his interpretation of the notation, and his use of it obscures some features of inferences undoubtedly made by the subjects which logicians call logical. This vitiates the claim that propositional logic provides the essential structure of the final stage of logical development. For the same reason it is uncertain that Piaget's model completely describes the difference between his stages II and III. Later, I shall raise questions concerning the relation of propositional logic and the operational schemata. The argument will require a somewhat technical discussion of Piaget's logic.

Propositional logic is often called 'truth-functional' because there are rules by which, given an assignment of truth values ('true' or 'false') to the sentence letters ('*p*', '*q*', etc.) of a given formula, a unique truth-value is assigned to a given formula; e.g. '*p* ∨ *q*' is true if '*p*' is true and '*q*' false, and false if '*p*' and '*q*' are both false. This is in accord with the interpretation of the letters as representing propositions which are true or false, and of the formulae as representing

† Inhelder, B. & Piaget, J. *The Growth of Logical Thinking from Childhood to Adolescence*. London: Routledge and Kegan Paul. 1958. Pp. xxvi + 356. 32s.

‡ The present article is the second of twin reviews of the Inhelder and Piaget book. The first review, 'A Psychologist's viewpoint' by Jerome S. Bruner, appeared in the last issue of this *Journal*.

propositions whose truth or falsity depends on that of the propositions represented by the component letters.

Piaget seems at first glance to use this interpretation of the logic. In the *Traité de logique* he says: 'Interpropositional operations leave unchanged the propositions p , q , etc., in order to envisage only their truth and falsity and to compose them among themselves as elements of a new system.' (p. 219). However, he does not mean by this quite what is meant by most logicians, and since he does not acknowledge the difference, his usage is confusing to a logician. Piaget usually uses the letters, in effect, to represent propositional functions, e.g. what is expressed by sentences containing free variables, such as ' x is a U.S. Senator', which are only true or false when values of the argument variables are specified, e.g. by replacing ' x ' by 'J. W. Fulbright'.

What corresponds for propositional functions to the truth value of a proposition is its extension, the class of values of the argument variable for which the resulting proposition is true. An extension of the propositional calculus, called the predicate calculus, the functional calculus, or quantification theory, contains formulae which have truth-values when extensions are assigned to predicate letters; e.g. ' $(\exists x) (Fx.Gx)$ ' is true if the extension assigned to ' F ' and that assigned to ' G ' have elements in common, otherwise false. (It can be read 'Something is both F and G '.)

Piaget explains the meaning of the truth-functional connectives by means of a special interpretation in which the letters ' p ', ' q ' represent propositional functions. We call this the 'official' interpretation, because although he suggests it is not the only one, it is the only one he explains at any length in TL (pp. 229 ff.). Given two letters ' p ' and ' q ', all the formulae containing no other letters fall into sixteen 'equivalence classes'; i.e. two formulae belonging to the same class if and only if they have the same truth-value for each assignment of truth-values to ' p ' and ' q '; we can say they express the same truth-function. Each class except the class of inconsistent formulae (false for every assignment) contains exactly one formula in complete disjunctive normal form (c.d.n.f.), i.e. such that it is an alternation of distinct conjunctions each of which contains each letter, with or without a negation sign. It is thus an alternation of ' pq ', ' $p\bar{q}$ ', ' $\bar{p}q$ ', ' $\bar{p}\bar{q}$ '.

Now let P and Q be the extensions of the propositional functions which ' p ' and ' q ' are taken to represent. Suppose our variables range over vertebrates. If ' p ' represents ' x is a vertebrate with lungs' and ' q ' represents ' x is a vertebrate with gills', then P = the class of vertebrates with lungs and Q = the class of vertebrate with gills. These two classes partition the range of values of the variables into four subclasses PQ , $P\bar{Q}$, $\bar{P}Q$, $\bar{P}\bar{Q}$, in an obvious correspondence with ' pq ', ' $p\bar{q}$ ', ' $\bar{p}q$ ', ' $\bar{p}\bar{q}$ ', i.e. PQ = the class of vertebrates with both lungs and gills, $P\bar{Q}$ = the class of vertebrates with lungs and no gills, etc. Then on the official interpretation, a formula $S(p, q)$ is true if and only if

(a) If p^*q^* , where $p^* = 'p'$ or $p^* = '\bar{p}'$, and $q^* = 'q'$ or $q^* = '\bar{q}'$, occurs as an alternand of the c.d.n.f. of $S(p, q)$, then the corresponding class P^*Q^* is non-empty.

(b) If p^*q^* does not so occur, then P^*Q^* is empty. For example, the formula ' $p \vee q$ ', whose c.d.n.f. is

$$pq \vee p\bar{q} \vee \bar{p}q \quad (1)$$

is true if and only if PQ , $P\bar{Q}$, $\bar{P}Q$ are non-empty, and $\bar{P}\bar{Q}$ is empty, i.e. if and only if

$$(\exists x) (x \in P.x \in Q). (\exists x) (x \in P.x \notin Q). (\exists x) (x \notin P.x \in Q). - (\exists x) (x \notin P.x \notin Q), \quad (2)$$

which is equivalent to

$$(x) (x \in P. \vee. x \in Q). (\exists x) (x \in P.x \in Q). (\exists x) (x \in P.x \notin Q). (\exists x) (x \notin P.x \in Q). \quad (3)$$

Thus ' $p \vee q$ ' is interpreted as the true statement: 'There are vertebrates with lungs and gills, there are vertebrates with lungs and no gills, there are vertebrates with gills and no lungs; but there are no vertebrates with neither gills nor lungs, i.e. all vertebrates have either gills or lungs.' If ' Px ' represents ' $x \in P$ ' and ' Qx ' represents ' $x \in Q$ ', Piaget establishes a correspondence between a consistent truth-functional formula such as ' $p \vee q$ ' and a quantificational formula, viz.

$$(\exists x) (Px.Qx). (\exists x) (Px. - Qx). (\exists x) (-Px.Qx). - (\exists x) (-Px. - Qx), \quad (4)$$

which we shall call the Piaget formula associated with ' $p \vee q$ '. Moreover, every possibility as to which of the classes PQ , $P\bar{Q}$, $\bar{P}Q$, and $\bar{P}\bar{Q}$ is non-empty can be expressed in this way by a truth-functional formula.

This interpretation is clearly what Piaget has in mind on p. 76 of GLT, where he speaks of the verification of the formula

$$pqx \vee p\bar{q}x \vee \bar{p}q\bar{x} \vee \bar{p}\bar{q}\bar{x}. \quad (5)$$

He means that the subject has found that all these combinations occur, and that only these combinations occur. More exactly, let ' Px ', ' Qx ', ' Xx ' represent the propositional functions represented by ' p ', ' q ', ' x ' respectively. Then (5) means the same as

$$(x) (Px.Qx.Xx. \vee Px. - Qx.Xx. \vee. - Px.Qx. - Xx. \vee. - Px. - Qx. - Xx). \\ (\exists x) (Px.Qx.Xx). (\exists x) (Px. - Qx.Xx). (\exists x) (- Px.Qx. - Xx). (\exists x) (- Px. - Qx. - Xx) \quad (6)$$

The interpretation is always followed in the case of tautological formulae (TL, pp. 229-30; GLT, pp. 76, 293-4).

Concerning this interpretation, certain things must be said. In the first place, it is not an interpretation in the sense in which that work is ordinarily used in logic, since the theorems (tautologies) of truth-functional logic are not always true under it. Tautology (called by Piaget 'complete affirmation') is true only in the case where all four classes are non-empty; it corresponds to the schema

$$(\exists x) (Px.Qx). (\exists x) (Px. - Qx). (\exists x) (- Px.Qx). (\exists x) (- Px. - Qx), \quad (7)$$

which is not a valid formula of quantification theory. Thus on p. 76 Piaget speaks of the subjects' verifying the tautology q^*x , which shows that weight has no effect, which is not logically necessary. Similarly certain relations of implication and negation no longer hold. At most one of the c.d.n.f.'s is true for a given choice of P and Q , for if a conjunct p^*q^* occurs in a c.d.n.f. S but not in a c.d.n.f. T , then by (a) if S is true, P^*Q^* is non-empty, but by (b) if T is true, P^*Q^* is empty. Then if a formula S implies a formula S' without being equivalent to S' , e.g. if $S = 'pq'$ and $S' = 'p \vee q'$, on Piaget's interpretation whenever S is true, S' will be false. Given a formula S , the Piaget formula obtained from the negation of S is *never* the negation of the Piaget formula obtained from S or equivalent to it.

The application of substitution yields anomalies. If ' $pq \vee \bar{p}\bar{q}$ ' is considered as a c.d.n.f., its Piaget formula is

$$(\exists x) (Px.Qx). (\exists x) (- Px. - Qx). - (\exists x) (Px. - Qx). - (\exists x) (- Px.Qx). \quad (8)$$

But if ' $Px.Qx$ ' is substituted for ' Px ' and ' $- Qx$ ' for ' Qx ' in (4), the first conjunct is replaced by ' $(\exists x) (Px.Qx. - Px. - Qx)$ ', which is inconsistent, so that the whole formula is inconsistent and therefore not equivalent to (8). One would also get an inconsistency if one substituted for ' Px ' and ' Qx ' the Piaget formulae of ' pq ' and ' $\bar{p}\bar{q}$ ' instead of ' $Px.Qx$ ' and ' $- Px. - Qx$ '. Thus if a truth-functional formula S' is obtained by substitution from a formula S , then the Piaget formula of S' is not necessarily equivalent to the formula obtained by the corresponding substitution from the Piaget formula of S , according to two natural meanings of the 'corresponding substitution'.

This fact shows that Piaget cannot have this interpretation in mind in all his use of truth-functional formulae. For the meaning of conjunction and alternation in c.d.n.f.'s like (1) and ' $pq \vee \bar{p}\bar{q}$ ' cannot be the same as the official one. The first conjunct of (3) and (6) suggests taking these operations truth-functionally, as does Piaget's appeal to well-known truth-functional equivalences. When he speaks of conjunctions or 'associations' as true, he means true *sometimes*. 'If the conjunctions $(p.q)$, $(\bar{p}.q)$, and $(\bar{p}.\bar{q})$ are true while $(p.\bar{q})$ is false, then there is implication' (TL, p. 233) means 'If $(\exists x) (x \in P.x \in Q)$, $(\exists x) (x \notin P.x \in Q)$, $(\exists x) (x \notin P.x \notin Q)$, while $-(\exists x) (x \in P.x \notin Q)$, then there is implication'. (If the statement were meant truth-functionally, it would be absurd.). When he says, 'Operation $p[q]$ is equivalent to $pq \vee \bar{p}\bar{q}$ and its negation $\bar{p}[q]$ to $\bar{p}q \vee p\bar{q}$. Thus, both operations amount to affirming (or denying) that p is true when q is true or false' (GLT, p. 301), he seems in view of the explanation on p. 302 to mean that ' $p[q]$ ' means ' $(\exists x) (x \in P.x \in Q). (\exists x) (x \in P.x \notin Q)$ '. (2)

He seems not to distinguish clearly between interpretations. After explaining the official interpretation of 'implication', (3) (TL, p. 233), he says that $p \supset q$ corresponds to class inclusion $P < Q$, i.e. to the emptiness of $P\bar{Q}$, but on the official interpretation it requires also the non-

emptiness of PQ , $\bar{P}Q$, $\bar{P}\bar{Q}$. (If ' $<$ ' is interpreted to signify proper inclusion, then it follows from $P < Q$ that $\bar{P}\bar{Q}$ is non-empty.) The interpretation as inclusion occurs on p. 17 of GLT, where he says that the subject understands and establishes the 'implication $p \supset q$ ', where ' p ' can be taken to represent ' x is the angle of incidence of a trial y ' and ' q ' to represent ' x is the angle of reflexion of a trial y '. (The subjects are to fire a ball so as to bounce off a wall and hit a certain object.) Then what Piaget means is that the propositional function represented by ' $p \supset q$ ' is always true, what might be symbolized as ' $(x)(y)(Pxy \supset Qxy)$ '. At the beginning of the general explanation (GLT, p. 217), he apparently has the official interpretation in mind, but on pp. 299-300 he says ' $p \supset q$ ' is compatible with ' $q \supset p$ ' (also pp. 299-300, p. 17; TL, p. 234), which is not true under the official interpretation; on pp. 17-18 he speaks of the subjects' verifying *both* these implications and of their conjunction as amounting to the equivalence of p and q .

One must protest against so much ambiguity and obscurity in the use of logical symbolism. As Piaget points out, the symbolism was invented for precision (TL, p. 17). It is not sufficient to say that the intention was to present symbolic logic 'intuitively' (TL, p. v); the first requirement of an intuitive presentation is that it should be clear.

It is puzzling that Piaget does not use quantification theory, since by means of it the ambiguities can be easily disentangled. (4) It must be understood that these considerations are not just of exposition but of logical and, as we shall see, psychological substance. Piaget's methods would make hash out of any attempt to give a formal criterion of logical implication, equivalence, or validity (cf. note (4)). Moreover, what of his claim that there are sixteen 'binary operations'? If he means that there are sixteen distinct truth-functions of two arguments, this is of course demonstrably true. If he means that there are only sixteen distinct logical functions of two propositional functions of one variable, this is demonstrably false, even if we assume that what is meant by 'logical functions' involves expressibility in the notation of truth-functions and quantification. In order to take in more than the sixteen Piaget formulae, he uses the notation ambiguously. It seems to me that there is no ground for supposing that a formula used in two such different interpretations expressed the same 'logical operation' in each case. Whatever such identity may mean, it certainly precludes that the formula should be true in one case and false in another, as can happen, for example, if ' $p \supset q$ ' is read now as ' $(x)(Px \supset Qx)$ ', now by the official interpretation. There is nothing against Piaget's saying that ' $p \vee q$ ' and ' $\bar{p}q \vee p\bar{q} \vee \bar{p}\bar{q}$ ' express the same logical operation, for these are logically equivalent, but it is false or senseless to say that ' $(x)(Px \vee Qx)$ ' and (4) express the same logical operation.

What fails to make logical sense can hardly make psychological sense in a study of intellectual development. The only psychological justification for speaking of 'the same operation' in cases of logical non-equivalence would be if the subjects were unaware of the non-equivalence or act as if it were irrelevant. No evidence is presented to show that this is true of Miss Inhelder's adolescents, and if it were it would surely prove that the development of the subjects' capabilities of logical reasoning is incomplete. In either case, Piaget's logic fails to describe a state of 'final equilibrium'. In fact, he assumes the subjects are capable of inferences which are not truth-functional or covered by a theory of transformations of Piaget formulae, e.g. from ' $(\exists x)(Px - Qx)$ ' to ' $-(x)(Px \supset Qx)$ ' (cf. p. 299). In order to verify experimentally statements of the form of Piaget formulae, subjects must appeal to more than their logical relations to other statements of the same form. Of course Piaget does not deny this, but his conclusions rest on a failure to analyze what is involved.

Although Piaget does not make it clear, the set of Piaget formulae does play an interesting role in the quantification theory of propositional functions of one argument. It can be proved that every formula of this theory containing only two (one-place) predicate letters ' P ' and ' Q ' is equivalent to a formula constructed by truth-functional connectives from the four formulae ' $(\exists x)(Px.Qx)$ ', ' $(\exists x)(Px.-Qx)$ ', ' $(\exists x)(-Px.Qx)$ ', and ' $(\exists x)(-Px.-Qx)$ '. (5) The resulting formula can be put into complete disjunctive normal form. Then each formula, unless it is in occurrence of each of the above four formulae, each of which is a conjunction containing exactly one alternation of Piaget formulae, negated or not. That is, each such formula is an alternation of Piaget formulae. (It follows that there are 2^4 equivalence classes of formulae containing just ' P ' and ' Q '.) This can be generalized to formulae with three or more predicate letters, provided they are one-place.

Let us call a statement obtained by substitution of predicates (e.g. ' x is an animal') for the letters of a quantificational formula a quantificational statement. If a formula is a Piaget

formula, we call the statement a Piaget statement. It follows from the last paragraph that in so far as the solution to a problem can be given in terms of a quantificational statement with two predicates, the best possible solution is in the form of a Piaget statement. For if it is known which of the sixteen Piaget statements holds, the truth of any quantificational statement in the two predicates can be mechanically determined, simply by computing the above-mentioned normal form which is an alternation of Piaget statements, and observing whether the verified Piaget statement is among the alternands. If it is, then clearly the statement in question is true. Otherwise, since at most one of the sixteen Piaget statements is true, all the alternands of the statement in question are false, and so is the statement itself.

This fact gives the Piaget formulae a central role in analyzing experiments such as Miss Inhelder's, where the factors can be sorted out and general laws arrived at by intelligent experimentation. The solution could be given in the form of a Piaget statement which would express the fact that the strongest possible law had been verified, and the other possibilities refuted. ('Complete affirmation' would express the fact that all possible laws had been refuted.) As we have shown, however, Piaget's analysis could not be put entirely in terms of Piaget formulae, and in fact the official interpretation is not so conspicuous in GLT as in TL.

The fact that the problems involve laws raises another question. It is an old philosophic problem to justify the assertion of general laws on the basis of observation of particular cases. It would certainly be too much to expect the authors to have the solution to this problem, but the effect of the notation is to slur over the fact that the subjects *do* verify laws, or rather to obscure the distinction between laws and particular cases, expressed in quantificational notation by the difference of ' $(x)Fx$ ' and ' Fa '.⁽⁶⁾ This is one of the reasons by Piaget's use of notation is ambiguous. It appears that the stage III subjects are clearer about what is involved in verifying a law than are the less advanced ones, but subjects are said to distinguish 'all' and 'some' at stage IIA (p. 30). Although some kind of 'generalizing tendency' must be implicit in the child's behaviour from the beginning, it seems that subjects of stage I have no idea how to justify or refute a law, and subjects of stage II only in special cases. That stage III subjects seem to look more consciously for a law suggests that they have a more explicit idea of universal quantification.

In the analysis of a number of problems, Piaget appeals to four transformations I, N, R, C on truth-functions (or formulae). I is simply the identity; N is negation. Given a formula S , the reciprocal $R(S)$ is the formula obtained from S by negating every letter. The correlative C is the negation of the reciprocal. If S and T are logically equivalent, then $I(S), N(S), R(S), C(S)$ are equivalent to $I(T), N(T), R(T), C(T)$ respectively, so that I, N, R, C can be construed as transformations on truth-functions. The four transformations form a group; i.e. combining two of them yields a third, and each has an inverse (namely itself). The group has the multiplication table

$$\begin{aligned} I^2 &= N^2 = R^2 = C^2 = I, \\ IR &= RI = R; IN = NI = N; IC = CI = C, \\ NR &= RN = C, \\ CR &= RC = N, \\ CN &= NC = R. \end{aligned}$$

This group is isomorphic to a well-known group called the four group.

Empirically, the four group has a number of realizations of the following type. Consider a system in which there are factors p_i and q_i such that p_i produced the same effect as q_i , and factors p_i^*, q_i^* which produce opposite effects to p_i, q_i respectively. Then let

$$\begin{aligned} A(p_i) &= p_i^*; A(q_i) = q_i^*; A(p_i^*) = p_i; A(q_i^*) = q_i; \\ B(p_i) &= q_i^*; B(q_i) = p_i^*; B(p_i^*) = q_i; B(q_i^*) = p_i. \end{aligned}$$

A can be called 'inversion', B 'reciprocity'. The group consists of I, A, B, AB .

A clear example is the projection of shadows by holding a disc between a point-source of light and a screen (chap. 13). Let p_n be the operation of increasing the distance of the disc from the source by a factor of n ; let q_n be the operation of replacing a disc of diameter d by one of diameter d/n . p_n^* is dividing the distance by n ; q_n^* is replacing a disc of diameter d by one of diameter nd . Then inversion and reciprocity are as in the last paragraph. p_n is compensated either by reversing it (p_n^*) or by increasing the diameter of the disc by the same factor (q_n^*).

The same structure is found in problems of mechanical equilibrium. Adding a weight to one arm of a balance is compensated either by removing the weight or by placing an equal weight on the other arm, at the same distance from the centre. If the distance of a weight from the centre is increased by a factor of n , this is compensated by dividing the weight by n , as well as by annulling the increase.

Any system which follows a proportionality law will exhibit this structure. In the case of the shadows, the size of the shadow depends on the ratio d/D , where d is the diameter of the disc. The above-described compositions follow immediately.

The structure of the balance problem is more complicated. The position of the arms is a function of the two products $W_1 D_1$, $W_2 D_2$ of weight on each arm with its distance from the fulcrum. A given transformation p_n (say, an increase in W_1 by a factor of n) can be compensated in four ways (' $\rightarrow$ ' can be read 'changed to'):

$$\begin{array}{ll} p_n^*: W_1 \rightarrow W_1/n, & p_n': W_2 \rightarrow nW_2, \\ q_n^*: D_1 \rightarrow D_1/n, & q_n': D_2 \rightarrow nD_2. \end{array}$$

It seems that the four group is too simple to describe this fully, but it describes the situation where two of the variables are fixed. The advanced subjects seem to assume at the outset that if W_2 and D_2 are fixed, then an increase in W_1 compensates a decrease in D_1 . (7)

Conversely, the existence of a double compensation is a strong clue to the existence of a proportion. If an increase in one measurable factor is compensated by an increase in another, then that the equilibrium condition is the preservation of a ratio (direct proportion) is a natural hypothesis; if an increase in one measurable factor is compensated by a decrease in another, then it is plausible that the equilibrium condition is the preservation of their product (inverse proportion). (The hypotheses of difference and sum, respectively, are also natural, and it is not so clear how the subjects eliminate it, or why it does not occur to them.)

Thus there are connexions between compensation, the four group, proportionality, and equilibrium which are simple and intrinsic to the structure of the systems themselves. Piaget undertakes to relate these to the realization of the four group by the transformations I , N , R , C on truth-functions. His thesis appears to be that the subjects' acquisition of propositional logic enables them to solve the problems *by way of a correspondence* between the logical group $INRC$ and the isomorphic group connected with the physical system. This connexion is not very clear and seems rather tenuous. I shall consider in detail the cause of the balance.

'The subjects understand the following relations of inversion and reciprocity (the $INRC$ group but with $p.q$ chosen as the identical operation I):

$I(p.q)$ = to increase simultaneously the weight and the distance on one of the arms;
 $N(\bar{p} \vee \bar{q}) = (p.\bar{q}) \vee (\bar{p}.q) \vee (\bar{p}.\bar{q})$ = to reduce the distance while increasing the weight or diminish the weight while increasing the distance or diminish both;
 $R(p'.q')$ compensates I by increasing both weight and distance on the other arm of the balance;
 $C(\bar{p}' \vee \bar{q}') = (p'.\bar{q}') \vee (\bar{p}'.q') \vee (\bar{p}'.\bar{q}')$ = cancels R in the same way that N cancels I . (8)

But, since $R(p'.q')$ is equivalent to compensating action $I(p.q)$ with a reaction (symmetry) on the other arm of the balance, we find that it can be written $\bar{p}.\bar{q}$; and since $(\bar{p} \vee \bar{q})$ is also equivalent to compensating the action N by symmetry, we can write it $(p \vee q)$. Therefore proposition (8) can be formulated as follows:

$$\begin{array}{ll} I(p.q), & \\ N(\bar{p} \vee \bar{q}) & \\ R(\bar{p}.\bar{q}), & (8a) \\ C(p \vee q)' & (p.178)(8) \end{array}$$

However, ' $\bar{p}.\bar{q}$ ' and ' $p'.q'$ ' (taking ' $-$ ' to be meant in the way it is elsewhere in the discussion of equilibrium, as we mean '*') (9) are closer to being in the negation-reciprocal relation in the above-explained empirical sense than are ' $\bar{p} \vee \bar{q}$ ' and ' $\bar{p}\bar{q}$ ', since $\bar{p}\bar{q}$ annuls pq , while $p'q'$ compensates pq without annulling it. $\bar{p} \vee \bar{q}$ only partly annuls pq if only one of $\bar{p}$ or $\bar{q}$ is performed.

Moreover, there is no evidence that $\bar{p} \vee \bar{q}$ or $\bar{p}' \vee \bar{q}'$ play any role in the actual reasoning of the subjects. The group which is more meaningful would be

$$I(pq) = pq.$$

$$N(pq) = \bar{p}\bar{q}.$$

$$R(pq) = p'q',$$

$$C(pq) = \bar{p}'\bar{q}',$$

but here the truth-functional relations of negation, reciprocity, and correlativity no longer hold.

The proportions (10) and (11) (p. 179) which Piaget connects with the subject's attack on the problem, appear to be derived from the proportions (9), the first of which is obtained from (8a). But (10) and (11) are not derived from the first of (9) at all and not from $Ix/Rx = Cx/Nx$ where 'R' and 'C' have their truth-functional sense, but only on the assumption $q = Rp$ or $q = Cp$. These are false truth-functionally, but the latter makes empirical sense because $\bar{q}$ compensates p . The relation of (12) to (11) is the same as that of (10) to (9) and the same consideration applies. It is only by analogy that these proportions are related to the truth-functional group. We could not take them literally in that sense, for otherwise we should have logical truths (10) used to schematize empirical facts.

The analogy would be close enough to be very illuminating if '—' and 'R' as used in analyzing the problem corresponded more closely in their logical behaviour to the same notations in logic. We start with the fact that a transformation p is compensated by two different transformations p^* and q^* . Hence we call p^* ' $\bar{p}$ ' and q^* ' Rp '. But already with conjunction, the parallel breaks down, for p^*q^* annuls pq , while $p^* \vee q^*$ (the candidate for ' $-(pq)$ ') does not.

In fact, Piaget seems to make a large assumption in apparently supposing that a subject who can use the system of binary truth-functions can, in some sense, use the truth-functional realizations of the four group. Reciprocity and correlativity are rather sophisticated relations of truth-functions, and it is hard to think of any behaviour which would show command of them which did not also show the subject's possession of an explicit concept of truth-function, i.e. that the subject knows some formal logic. I shall simply assert what I claim to have shown for the case of the balance, that the occurrence of the truth-functional form of the group in the subjects' thinking is nowhere convincingly shown.

We are thus left without a satisfactory explanation of the subjects' acquisition of proportions. If the question is simply why they are not understood earlier, then the answer may be simply that the relation of proportionality is too abstract to be discovered by enumeration of correspondences and therefore requires a hypothetico-deductive approach. But there remains the question why they come to this hypothesis so readily. It certainly seems to be by way of the group of compensation relations. Piaget seems to find the mere isomorphism of the group to the logical INRC group extremely suggestive.

This brings us to a somewhat obscure aspect of Piaget's thought. In some sense, what is crucial about the operations at the concrete and formal level is their algebraic structure, of groupings in the case of concrete operations, and of Boolean algebra, with the superimposed four group, in the case of formal operations. Thus in a later work Piaget writes: 'One can describe these new operations in using different more or less adequate formal languages (according as the isomorphisms obtained are more or less complete). One can consider them as a completion of the logic of classes. . . . One can express them in the language of the first-order functional calculus. . . or simply in that of the propositional calculus. For simplification, we shall adopt the latter notation by convention. What is important from the psychological point of view is anyway not the language chosen (for the logic of the real subject is isomorphic to an algebra as calculus more than to a formalized language) but the fact of the appearance of new structures, the lattice and the group of four transformations, which widely exceed the elementary groupings of the preceding level'. (*Logique et équilibre*, Paris, 1956, p. 97. It seems from this that Piaget may no longer hold the views of TL discussed in note 4.) Here Piaget seems to be saying that the main point is that these algebraic structures represent an expansion of the basic cognitive capacities of the subjects. This I have not disputed. But questions still arise. One can associate with the first order functional calculus a much more complicated algebra than general Boolean algebra. (P. R. Halmos, 'The basic concepts of algebraic logic', *Am. Math. Monthly*, 1956, 63, 363–87.) Why is this richer structure not included in the 'algebra as calculus' to which 'the logic of the

real subject is isomorphic'? In the case of the occurrences in the problems of the four group, if simply the fact that it is the four group is crucial, then the connexion with propositional logic which we claim Piaget has not made would in any case be superfluous. If we rest with the fact that the four group is part of the structure of the system, then is the isomorphism of one's 'logic' to an 'algebra' more than simply an ability to solve problems where that algebra is involved in an essential way? This might open the way to ridiculous extensions of the scope of logic: for Miss Inhelder's adolescents, it would include the rational numbers and some extensions of them; for a professional abstract mathematician, it would include an enormous and indeterminate variety of algebras.

The problem of characterising a stage of development by formal logic meets a well-known philosophical problem, that of determining the place of formal logic in the structure of knowledge. Actual inference can involve logic, relations of meaning, empirical fact, and mathematics. We do not know very clearly what is the principle of the separation of these four aspects. It is only with respect to problems of a clear-cut and simple structure that the formal element can be easily isolated.

To characterize a stage by algebra is at another remove of abstraction and raises more problems. For example, what would we say of subjects who solve problems involving more complicated groups than the four group, even of the same general nature? Indeed, as we suggested, we could describe the balance as involving such a group by virtue of the fact that each transformation has four compensating ones instead of two. The solution to the problem perhaps requires operations only in subgroups of this group which are isomorphic to the four group. But this is, on the face of it, a lucky accident. I do not claim that this use of algebra is wrongheaded in principle, but I do not know exactly what the criteria of adequacy are or how far the fact that the algebra is associated with a system of logic is to be pressed.

In general, it seems that Piaget's theory leaves us unclear as to what to expect of the performance of subjects on problems quite different from those discussed, or of subjects quite different from Geneva school-children. If it is claimed that the analysis by propositional logic 'goes more deeply into the heart of intelligence (*dans l'intimité des mécanismes intellectuels*) than arithmetical or statistical calculations... reaches the structures themselves—i.e. the operational mechanism as such and not simply its results or its more or less general conditions' (p. 270), then it must be said that only a very general light is shed on the operation of these structures in cases other than mathematics or experimentation. The observations on this subject in the last chapter are interesting but would not serve to explain details.

Notes

(1) Piaget says, 'Concrete structures or "groupings" of classes and relations derive from either inversion (classes) or reciprocity (relations), but entail no general synthesis of the two forms of reversibility' (p. 273). The reciprocal of a relation is its converse (p. 274). Piaget can hardly mean to say that at the concrete level negative relational statements are never made, but rather that this negation does not give rise to a system of operations.

(2) It seems that the statement

For example, if $p[q]$ and $q[p]$ are both true at once and alone true, one has $pq \vee p\bar{q} \vee \bar{p}q$, that is to say $p \vee q$. If $p[q]$ and $\bar{q}[p]$ are true at once and alone true, one has $pq \vee p\bar{q} \vee \bar{p}\bar{q}$, that is to say $q \supset p$. If $\bar{p}[q]$ and $q[p]$ are both and alone true, one has $pq \vee \bar{p}q \vee \bar{p}\bar{q}$, therefore $p \supset q$, and if $\bar{p}[q]$ and $\bar{q}[p]$ are both and alone true, one has $p\bar{q} \vee \bar{p}q \vee \bar{p}\bar{q}$, therefore $p | q$. (p. 267). makes sense only if we take the interpretation above sketched for ' $p[q]$ ', ' $q[p]$ ', ' $\bar{p}[q]$ ', ' $\bar{q}[p]$ ', and the official interpretation for ' $p \vee q$ ', ' $q \supset p$ ', ' $p \supset q$ ', and ' $p | q$ '. I do not find evidence that Piaget intended in GLT the official interpretation for the first of these groups. But it seems clearly to be meant in the explanations in TL, pp. 237–8. The corresponding Piaget formulae would be of little interest in the experiments of GLT, where ' p ' and ' q ' represent factors which can now be present, now absent according to what the experimenter does, so that P , Q , $\bar{P}$, $\bar{Q}$ are invariably non-empty. The interpretation of GLT however, does have an application to the independence of factors.

(3) Piaget's writings lack the important distinction between the truth-functional connective ' $\supset$ ', which following Quine (*Methods of Logic*, New York, 1950, §3) I should prefer to call the conditional, and the relation of implication between formulae (or statements) such that S implies T if the conditional with S as antecedent and T as consequent is logically valid. (*Ibid.* §7). Such

implications are appealed to in TL, but in spite of the presence of a chapter on the foundations of deduction, implication in this sense is never discussed.

(4) In fact, Piaget apparently thinks that this notation is superfluous. In discussing the theory of the syllogism he says: 'What the universal affirmative [Every X is Y] expresses is simply that the attribute Y is of extension equal to or superior to that of the subject and that the latter is taken in its whole extension. There exists a relation of the same type when one posits as true the implication ($p \supset q$). In fact, whatever may be the nature of the propositions p and q (that is to say that p and q belong to the categories A , E , I , or O), the implication ($p \supset q$) excludes the possibility ($p \bar{q}$), therefore "every time" that, in a relation ($p \supset q$), p is true, q is true. To the proposition p corresponds thus a class of arguments taken in its whole extension; otherwise one would not exclude ($p \cdot \bar{q}$)' (TL, pp. 363-4). What is suggested here and at the end of the paragraph is the interpretation of ' $p \supset q$ ' as ' $(x) (Px \supset Qx)$ ', although in between he gives an example where p and q are closed propositions, and ' $p \supset q$ ' is taken to mean: 'Every choice of extensions of the predicates such that p is true makes q true.'

In the discussion of the universal negative, ' $p|q$ ' is read ' $(x) (Px|Qx)$ ', but to show that propositional logic encompasses existential quantification, a quite different argument is used: 'In fact ($p \vee q$) expresses by definition the possible combinations ($p \cdot q$) $\vee$ ($p \cdot \bar{q}$) $\vee$ ($\bar{p} \cdot q$), which include ($p \cdot q$) and ($p \cdot \bar{q}$), therefore precisely the propositions corresponding to the classes PQ and $P\bar{Q}$. To take an example... if p = " x is an object" and q = " x is beautiful", the trilemma ($p \vee q$) signifies that either some objects are beautiful, or some objects are not beautiful, or some beautiful realities are not objects. In fact, if one did not invoke the quantification 'some', the two full realities are not objects. In fact, if one did not invoke the quantification ($p \vee q$), but to a dilemma propositions p and q would not give place to a non-exclusive disjunction ($p \vee q$), but to a dilemma or exclusive disjunction ($p \vee q$), that is to say $p = \bar{q}$ or $q = \bar{p}$. The trilemma presupposes therefore, in so far as it is a trilemma ($p \vee q$), the quantification "some", or what comes to the same thing, "sometimes true", (TL, p. 365). It is difficult to make sense of this passage; among other things it contains another interpretation still of ' $p \vee q$ '. If we regard this as a slip and assume Piaget really intends the official interpretation (since certainly his conclusions cannot be supported by arbitrarily spinning forth interpretations of the connectives), i.e. with ' Px ' for ' x is an object' and ' Qx ' for ' x is beautiful':

$$(x) (Px \cdot Qx \cdot \vee \cdot Px \cdot -Qx \cdot \vee \cdot -Px \cdot Qx) \cdot (\exists x) (Px \cdot Qx) \cdot (\exists x) (Px \cdot -Qx) \cdot (\exists x) (-Px \cdot Qx)$$

then the point seems to be that the trilemma requires that both

$$(\exists x) (Px \cdot Qx) \quad \text{and} \quad (\exists x) (Px \cdot -Qx)$$

be true (while ' $p \vee q$ ' excludes the former), but the former excludes the universal negative (No objects are beautiful), while the latter excludes the universal affirmative (All objects are beautiful). Moreover, only if both these formulae are true can ' $(x) (Px \vee Qx)$ ' be true while ' $(x) (Px \vee Qx)$ ' be false.

What is shown in these arguments, then, is not that the notation of propositional logic, as logicians define it, suffices to express quantification, or that a formal system using only truth-functional notation could be constructed in which what are ordinarily taken as quantificational, non-truth-functional inferences (e.g. syllogisms) could be formalized. What is shown is that on a certain special reading of certain truth functional formulae, they imply categorical statements of the theory of the syllogism. Thus Piaget's thesis that propositional logic involves quantification means no more than that quantification has been read into formulae which do not involve it explicitly, with the resulting above-described ambiguities in the use of the symbols.

On pp. 368-9, syllogisms are represented by truth-functional formulae. Apparently a formula such as

$$p \supset r \cdot pq \supset pr, \tag{A}$$

is to be read

$$(x) (Qx \supset Rx) \cdot (\exists x) (Px \cdot Qx) \cdot \supset (\exists x) (Px \cdot Rx).$$

The inadequacy of (A) as a means of representing the inference is shown by the fact that if the quantifiers were put on differently, the resulting formula would not be valid, e.g.

$$(\exists x) (Px \supset Qx) \cdot (x) (Px \cdot Rx) \cdot \supset (x) (Px \cdot Qx).$$

Moreover, certain of the truth-functional formulae given are not valid. (These correspond to syllogisms, which, according to the standard way of putting categorical statements into the

notation of modern logic, require an additional premiss of the form ' $(\exists x)Px$ '. One must suppose that this translation does not accurately represent the traditional theory. Cf. J. Lukasiewicz, *Aristotle's Syllogistic* (2nd ed., Oxford, 1957) and P. F. Strawson, *Introduction to Logical Theory* (London, 1952), chap. 6.

There is no argument in TL to show that the logical distinctions which make quantification necessary do not exist, but only a refusal to incorporate them into his notation. Moreover, there is no hint as to how the questionable procedure of dealing with the syllogism is to be generalized to the quantification theory of many-place predicates (' $x < y$ ', ' x prefers y to z ', etc.).

(5) Quine, *op. cit.*, §19. To get our result, Quine's reduction procedure must be modified slightly. Non-uniform schemata (containing variables other than ' x ') must be reduced to uniform ones by the procedure of §32. If a schema reduces to ' $\perp$ ' at step 3, it should be replaced by ' $(\exists x)(Px.Qx) \cdot -(\exists x)(Px.Qx)$ '. After step 4, components of the form ' $\exists F$ ' should be replaced by ' $\exists(FG) \vee \exists(F\bar{G})$ ', and similarly for ' $\exists \bar{F}$ ', ' $\exists G$ ', ' $\exists \bar{G}$ '.

(6) There is a further distinction between generalizations (universal quantifications) and a more restricted class said to be *lawlike*, which some contemporary philosophers are trying to clarify. A lawlike generalization carries some kind of natural necessity and allows the inference to a contrary-to-fact conditional. See Nelson Goodman, *Fact, Fiction and Forecast* (London, 1955). Since Piaget regards expression by a contrary-to-fact conditional as a linguistic mark of a 'true implication' (p. 279), he seems to intend true implications to be lawlike.

(7) Piaget says (p. 180) that they 'usually' begin by assuming ' $pq = R(pq)$ ', i.e., an increase in W_1 and a decrease in D_1 compensate the corresponding decrease in W_1 and the corresponding increase in D_1 . (His reading, 'Increasing the weight and reducing the distance on one of the arms is the same as reducing the weight and increasing the distance on the other arm', depends on the equivalence in effects of p' and $\bar{p}$ (our p^*) and of q' and $\bar{q}$.) I do not see the evidence for this. The subjects seem to suppose and then verify that an increase in W_1 compensates multiplicatively a decrease in D_1 . What this means is that the state of the balance (for fixed W_2, D_2) is a function of the product $W_1 D_1$. What Piaget says they assume (that is, the formula) follows.

(8) The notation is confusing here because an expression like ' $R(\bar{p}.\bar{q})$ ', which would normally mean 'the value of R for the argument $\bar{p}.\bar{q}$ ', means in fact 'the value of R for a certain argument, namely $\bar{p}.\bar{q}$ '. The normal use is found in equations like (12) and (15) on pp. 179, 180. Piaget does not bring out clearly the distinction between I, N, R , and C themselves, which would naturally be thought of as transformations *on* transformations of the system, and the set of four transformations *of the system* which are the values of I, N, R, C for a single argument.

(9) We have used a different symbol because in this context ' $-$ ' means for Piaget not the negation of a transformation, i.e. its not being performed, but the transformation which annuls it.

(10) A direct logical proportion ' $p/q = r/s$ ' can be construed as a truth-functional schema equivalent to ' $ps \equiv qr : p \vee s \equiv q \vee r$ '. Then $p/q = \bar{q}/\bar{p}$ is a tautology. A reciprocal proportion, $A/B = R(C/D)$, which is apparently to hold if $AD = R(BC)$ and $A \vee D = R(B \vee C)$ (in the use of Piaget, *Essai sur les transformations des opérations logiques* (Paris, 1952), appendix I, a proportion RR rather than a proportion R), can only be thought of as a relation between truth-functions and not as a schema, for if S and T are schemata, the relation $S = R(T)$ is invariant under logical equivalence, but not under material equivalence for an arbitrary assignment of truth-values. In this sense, " p " = $R("q")$ is false, but ' $p/q = R(\bar{p}/\bar{q})$ ' is (logically) true because ' $p\bar{q}$ ' = $R(\bar{p}q)$ and ' $p \vee \bar{q}$ ' = $R(\bar{p} \vee q)$.

PUBLICATIONS RECENTLY RECEIVED

Psychology: A study of a Science. Ed. by SIGMUND KOCH. Volume 1. *Sensory, Perceptual and Physiological Formulations.* London: McGraw-Hill. 1959. Pp. x+706. 77s. 6d.

This is the first volume of a major undertaking sponsored by the American Psychological Association, and subsidized by the National Science Foundation, as a mid-Century appraisal of current theories in psychology. It is to comprise seven large volumes to include learning, perception, social psychology and methodology; it is to be written by a total of some eighty specialists under the general editorship of Prof. Koch.

The volume to be reviewed here comes under *Study I: Conceptual and systematic*, and is entitled *Sensory, Perceptual, and Physiological Formulations*. (Study II is to be *Empirical Substructure and Relations with Other Sciences*). This first volume contains eleven contributions, apart from a preface by Dael Wolffe, who served as Chairman of a panel of consultants, and an introduction by Sigmund Koch. All the contributors are distinguished authorities writing in their own special fields.

The plan of the book is for each contributor to try to answer certain general questions posed by an advisory panel. Each author was asked to describe, for example: (a) 'the background factors which have influenced objectives, methods, and content of system; (b) orienting attitudes which have determined systematic objectives, methods, and conceptual content'. Then they were asked to describe the structure of the system as thus far developed, including: '(a) exhaustive itemization of *systematic* intervening, and dependent variables; (b) mode of definition of representative variables of each category; (c) major interrelations among constructs'. Several other general topics are suggested, such as: 'construction forms', 'mensuration and quantificational procedures', 'scope or range of application of system', 'history of the system to date in mediating research and evidence for the system.'

Each author was presented with the same list of topics and questions, with explanations of what was meant by these questions, and was asked to write his contribution according to the same formal scheme. How successful has this plan proved to be? It seems to have had very different effects upon the various contributors—if we may justly compare their performances here with their performances in books and papers of their own choosing.

We find first, J. C. R. Licklider describing not one but 'Three Auditory Theories'. As always, his writing is exciting: we feel his breath on the back of our neck. True, his spontaneity and sense of the topical is sometimes carried to the extreme—as when he tells us that Green has borrowed Marill's apparatus—but his contribution is brilliantly successful in conveying new and important ideas on signal detection and the recognition of speech. Perhaps he minimizes the debt which signal detection theory owes to communication and television engineers, and to physiologists and psychologists working on vision (in this country as well as in America), but it is a clear concise account of ideas which have hardly reached the text-books, and which are likely to prove of great importance wherever neutral activity is considered. Licklider's three theories are all concerned with the status of sensory thresholds: it is remarkable how alive this issue is when we consider that thresholds have been studied since the birth of modern psychology, and proverbially form the subject of our most boring practical class experiments.

C. H. Graham has undertaken the mighty task of describing 'Colour Theory'. This contribution is adapted from a forthcoming book and very little notice has been taken of the officially required form of the articles—'The Scheme' as we may call it. This author gives his own outline plan, and states, that 'some factual content of colour vision must be described before its topical form can be discussed... the present account will offer more than the usual amount of factual description and possibly less "methodological" discussion than might seem desirable to some psychologists'. He goes on to give an extremely condensed outline, in 145 pages, of recent colour theory; largely following Y. Le Grand and W. D. Wright. Perhaps the task was impossible in the available space; at any rate the result is too like the inscription on Newton's tomb in Westminster Abbey; in either case, if the reader can understand it is hardly necessary. We may look forward to his book.

Pirenne and Marriott, the only contributors from this country, write on the 'Quantum Theory of Light and the Psycho-Physiology of Vision'. They also adopt their own scheme. As always with these authors, their account is clear and accurate. It describes the now classical method, due to Hecht, Schlear and Pirenne, for estimating the number of quanta required for detection of a flash or light by comparing frequency-of-seeing curves with Poisson sum curves for various values of n . It also describes later work. Reference is made to such early workers as Rose, who was probably the first seriously to consider sensory discrimination as a matter of signal/noise ratio. Pirenne and Marriott are more interested, however, in random fluctuations of threshold due to the physical nature of light than to fluctuations due to random changes in the physiological system. Their approach is thus limited to vision, and it is primarily considered with the retina as a transformer of photic into neural energy rather than with the more general decision properties of the system, which suggest an information rather than an energy analysis. Their work is important in establishing the limits of the primary energy detector, but the more general implications of, in particular, H. B. Barlow's work might have received more attention. These authors did not adopt The Scheme because '(the article) can be considered as belonging to biophysics as much as to psychology'.

K. Ogle writes on 'Theory of Stereoscopic Vision'. He has made a serious attempt to organize his contribution in the expected manner. The result may be compared with his well-known book (*Researches in Binocular Vision*, 1950) on the same subject. Under the heading 'Formal Organization of the System', we find: 'The theory of stereoscopic vision lends itself to a rather explicit step-by-step formulation'. A table is then presented which 'relates the axioms and successive theoretical notions to the subjective experience and to the psycho-physiological-geometrical experimental procedures'. Under 'Axioms and Theoretical Notions' we find, for example, 'Hypothesis of retinal local signs', 'Panum's areas', 'stability of corresponding retinal points', 'Myosensory factors', 'the Horopter'. Under 'Subjective Experience' we find 'physiologically double images', and 'equalities of subjective depth'. In short, we find a list of necessary conditions for depth perception, some generalizations such as the horopter. But why this should be thought to be a table relating theoretical to experimental elements is a mystery. What Dr Ogle in fact does is simply to describe various empirical findings, many of great interest. He wishes for 'insight into causal relationships in terms of anatomical structure of physiological events... Failing (this), hypothetical constructs or intervening variables based on antecedent knowns are desirable, provided they are in accordance with the facts and still better, are testable'. He gives us several facts and poses interesting questions, but he attempts no formal theory in spite of The Scheme.

'Lunenburg's Theory of Space Perception', in which visual space is regarded as non-euclidean, is discussed by Albert A. Blank. The theory is essentially mathematical, with empirical features logically analogous to those of a geometry. The discussion fits well into The Scheme.

'The Systematic Psychology of Wolfgang Köhler' is discussed by W. C. H. Prentice. He finds it 'hard to epitomise and analyse Köhler's views for this series', partly because 'in the rather special sense of the present volume, the 'system' is probably not a system at all... Still (Köhler's) viewpoint is a consistent and an important one'. Perhaps it is fair to say that the more clearly the Gestalt position is expressed the more unsatisfactory, and even silly, it appears. Or do the following make some sort of sense? 'Should we not try to write useful laws about the experienced phenomena of order instead of trying to explain them away as instances of judgement, "learning" or other secondary events?' Now what does this mean? In particular, what is it to 'explain away' something? What is lost except mistakes? To take another example: 'A common misunderstanding of the Gestalt-theoretical position has been the interpretation that it is antianalytic, whereas it merely insists that the type of analysis chosen should reflect the realities of the phenomena investigated'. Unfortunately it would take too long to attempt an analysis of this sentence; it would be an interesting exercise. The Gestalt doctrine of isomorphism is taken seriously by Prentice, as 'a working hypothesis that gives structure to the total theory... It makes possible the general assumption that the structural properties of brain events must be something like the properties of experience.' Quite absent are experimental studies, such as those of the late K. S. Lashley, which purport to give evidence against isomorphism. Even obvious general objections are ignored. For example, it is absurd to suppose that the brain can physically become like such things as perceived colours or tunes, even if it might have 'traces' shaped like

perceived circles or traction engines. Model traction engines in the brain possibly, but *blue* traction engines in the brain? Yet, if *shape* can be isomorphically represented should not *colours*, sounds and smells be thus represented if the theory is to be sufficiently general? Further, to represent shape by isomorphic shapes, (pictures or models) would be wasteful of storage space and highly inconvenient for transfer and computing. What is *done* with the isomorphic shapes? Prentice's discussion fails even to raise the difficulties.

One of the most successful contributions is J. J. Gibson's 'Perception as a Function of Stimulation'. He has kept largely to The Scheme, and he has written a stimulating account of his theory, which is that for every sensory experience there is an unique set of sensory conditions which could in principle be described. Without the aid of any figures, (which he uses to great effect in his book *The Perception of the Visual World*, 1950) he presents a clear, and to a great extent compelling, account of perception with an explicit programme for research. As it stands, it is a purely psychological theory: 'the bare logic of the theory is extremely simple: there are dependent variables of perception and there are independent variables of stimulation. No intervening variables are presumed. Do the mathematical properties of the flowing array of stimulus energies really specify all the known properties of the phenomenal world of the observer? The writer ventures to suggest they do.' Thus the brain, *et al.*, is left out. Perhaps Prof. Gibson might feel able to include it when he has decided what it has to do.

L. Postman and E. C. Tolman give us an account of 'Brunswik's Probabilistic Functionalism'. They follow The Scheme closely, and they would probably accept the implied philosophy of such a scheme, for they write: 'Two major problems have faced psychology in its struggle for scientific maturity: the development of an objective methodology and the construction of a conceptual framework appropriate to the complexities and subtleties of the behaviour of organisms ... It was Brunswik's thesis that they could be reconciled in a psychology which (a) rested upon the methodological foundations of the physical sciences, and (b) focused its analysis on the achievements of the organisms in adjusting to the environments.' These authors proceed to give a most useful summary of Brunswik's ideas, with special emphasis upon the role of probability ideas: 'It is essential to recognize that probability laws may be just as objective and precise as strict laws. Whether one or the other applies is a question of empirical fact.'

'Adaptation level Theory' is discussed, as one would expect, by H. Helson. His first words are: 'Although this presentation is not cast explicitly in the form suggested by the editor, it has profited greatly from his outline and logical analysis of models'. Helson covers a wide range of topics with clarity, and has interesting things to say of each. He hints of a closer tie up with recent servo theory, which might make many of these ideas more precise and more amenable to experimental treatment.

Together with the first contribution, the most lively is D. O. Hebb's 'A Neuropsychological Theory'. He writes with charm and candour of the development of his theory. Thus: 'By the time the book *The Organisation of Behaviour*, 1949 was being written it was clear that the theory could not be as satisfactory as at times in the construction it had promised to be, and some protective manoeuvring had no doubt begun.' Of his early academic record: 'B. A. Dalhousie, 1925, with a distinguished record (lowest average without actually failing) ...' Hebb largely rejects The Scheme, but we are given lively comments on profound matters. He regards the limits of prediction as set by our intellectual capacity: 'it seems to me that the scientific method is to make the working assumption of determinacy up to the point at which it is conclusively refuted ...'. Could it be that those theorists who look for underlying casual processes would tend to adopt this view, while those content with hypothetical postulates at the behavioural level may prefer to invoke indeterminacy as a principle? Of especial interest is Hebb's account of why he regarded the Senden data on early blindness as important. This contribution is likely to be read and quoted for a long time.

The final contribution is C. T. Morgan's 'Physiological Theory of Drive'. Largely following The Scheme, he gives a too short account of his views. This topic will, no doubt, be discussed at greater length in the next volume.

Having indicated what the book contains, we may note what is absent. The editor explains that some invited authors fell by the wayside; but were they quite irreplaceable? There is a really serious absence of physiological, cybernetic or engineering ideas in this book. No regard is paid to anatomical structure, or to the efficiency of the sensory systems—apart from simple thresholds—or to such topics of contemporary interest as pattern recognition. A statement of the

berneticist's position, let us say by Warren McCulloch, would seem to be essential here. This type of thought and experiment is surely more important than the Lunenburg theory.

There is no sustained attempt to defend any position against criticism. Also, there is no attempt to answer or to resolve difficulties seen by the clash of theories. Given that The Scheme as deemed necessary (and it has clearly been used with the greatest tact and discretion) it seems unfortunate that it did not serve to bring into the open issues of live controversy. One's impression is that members of each persuasion are but little influenced by the views or the motivations of others. Certainly the gloves are not taken off here.

The book is a series of accounts of theories. One is used to this in the history of science—and perhaps this is contemporary history—but it seems more likely that what is important at present is not the formulation of theories according to set schemes, so much as indication of strategies for research. If undue emphasis is placed on formal theories before successful research strategies have been evolved, there is a clear danger that experimental results may be regarded as evidence for or against rival theories (and nearly all of them are certain to be wrong) instead of letting the data suggest new experiments and new research strategies. No doubt theories can be important, and may even be thought of as the output of research, but there is an emphasis on theories as closed systems, which is dull and self protective, when they can be bases for discussing crucial issues, and aids to the development of research programmes, which is important and exciting. The self-protective character of most of the theories discussed may be due to the fact that they are of the 'purely psychological' kind. Only Hebb seriously suggests a search for causal mechanisms. Surely explanation of sensory processes must be in terms of what goes on in the box of tricks. We do not understand a trick until we know what is in the box and how it works. A conjuror would not hesitate to take advice from engineers who know about other mysterious boxes (even if such advisors would not be granted membership of the Magic Circle) and he would be almost certain to learn some useful new tricks.

R. L. GREGORY

Psychology: A study of a Science. Ed. by SIGMUND KOCH. Volume 2. *General Systematic Formulations, Learning, and Special Processes*. London: McGraw-Hill. 1959. Pp. x + 706. 77s. 6d.

This is the second of seven volumes planned by a committee of the American Psychological Association and made possible by funds granted by the National Science Foundation. The aim is to take stock of developments in the many different aspects of Psychology. In view of the amount and diversity of empirical findings and the inevitable pre-occupation of psychologists with theoretical and methodological issues, it would appear perfectly justifiable to employ approximately eighty authors with adequate consulting services and ask each to treat his special topic early and candidly.

In the main this is what the fourteen authors of Volume 2 have done. Dael Wolfe in the Preface briefly describes the background of the project. Sigmund Koch (editor) indicates the place of Volume 2 in the series. Dorwin Cartwright discusses how extensions from Lewin's concepts and the empirical studies which he inspired may be referred to in further investigations. E. C. Tolman in his paper, 'Principles of Purposive Behaviour', begins very diffidently and concludes delightfully by telling the reader, who may discern points of weakness, that E.C.T. doesn't much care. A serious reader, however, could afford to overlook this paper and when, much later in the book, he reads Harry Harlow's 'Learning Set and Error Factor Theory', discussing the important findings from the Wisconsin Primate Laboratory, he may find himself turning back to Tolman's paper. Learning theory in fact receives a great deal of attention. E. R. Guthrie examines the implications of 'Association by Contiguity'. Neal Miller, writing on 'Liberalisation of Basic R Concepts', indicates how these concepts may serve as a basis of orientation for a wide range of behaviour in many species. F. A. Logan compares 'The Hull-Spence Approach'. W. K. Estes treats many aspects of 'The Statistical Approach to Learning Theory'. A. L. Irion presents a good case for renewed attention to the somewhat neglected field of 'Rote Learning' and the impact of technology is reflected in the papers on 'Information Theory' by F. C. Frick, and 'Linear Frequency Theory as Behaviour Theory' by D. G. Ellison. R. A. Hinde gives a succinct classification of the rapidly growing literature on 'Some Recent Trends in Ethology' and half-

way through the book B. F. Skinner in 'A Case History in Scientific Method' emphasizes that the curiosity of the individual scientist, unhindered by formalism, is an important factor in scientific progress.

The work is well arranged, although the individual contributions differ appreciably in conception. Every chapter provides a long and useful bibliography and in all cases the treatment is orientated towards the future as well as the past. It is a matter for debate as to whether or not a shorter work, confined to results which have been confirmed and a much condensed statement of issues, would not have been more useful; but there can be no doubt that the present volume will be a very useful work of reference to the advanced student of Psychology and does give a good and stimulating account of an important sector of psychological theory at this time. The facts are that compared with *Psychologies* of 1925 and *Psychologies* of 1930, there are now many more findings and possibilities to provoke attempts at co-ordination and surely this is evidence of health and vigour; but all of us who have sinned can reflect that in the current resources of Psychology, there is no dearth of words.

F. V. SMITH

Motivation: A systematic Reinterpretation. By DALBIR BINDRA. New York: The Ronald Press. 1959. Pp. vii + 361. \$5.50.

This is not a text-book on motivation, although in its breadth and detail of treatment it may be of some value in this respect. The limitations of the work for this purpose arise from the way the material is organized and from the restricted, rather superficial treatment of the more physiological aspects of the subject from which stems much of the present interest in this field.

It is Dr Bindra's intention to provide an account of what he calls 'motivational activities' which will bring these aspects of behaviour into some 'systematic' relation with such extensively studied phenomena as learning, discrimination, etc. As the author acknowledges, the method of analysis is closely related to those of Skinner and of Hebb. The use of 'hypothetical constructs' is avoided as far as possible in favour of an attempt to classify, in an orderly fashion, the factors which control the development and performance of 'motivational activities'. This is the contribution of Skinner. Questions concerning the mechanisms involved are regarded as being of a physiological nature, and here Hebb's physiological ideas are presented and further developed.

In so far as there is a single fundamental thesis in the book, it is that motivational activities are learned according to the same principles of stimulus-response conjunction, reinforcement, etc., which govern any other instances of instrumental learning. Dr Bindra argues, in addition, first that no previous account of the nature of reinforcement is uniquely correct—e.g. drive reduction, and consummatory response can both be shown to be efficient reinforcing agents. Secondly, it is maintained that no 'end-determining system' is required to explain the goal-seeking nature of motivated behaviour. While the first of these two subsidiary arguments might lead others to seek another class of events, other than drive reduction and consummatory responses, which might reconcile the conflicting evidence and prove to be 'the' reinforcing factor, Dr Bindra is true to the tradition of Skinner and is content if a list can be provided of the factors which may be observed to influence probability of response. Nevertheless, it is strange that the possibility that reinforcement may always consist in the reception of certain crucial stimuli is never considered. The implications of the second subsidiary argument are not clear. It is accepted that a principal descriptive characteristic of motivated behaviour is its goal-seeking quality, although there appears to be some inconsistency between different parts of the book as to whether this is a *definite* characteristic of such behaviour. However this may be, it seems to be suggested that we should not search for specific goal-setting and recording systems. The nature of reinforcement remains an unsolved physiological question, while the elicitation of motivated behaviour is discussed in terms of the concept of 'level of arousal' and with some reference to factors arising from the study of blood-chemistry. It is difficult to determine, from the text, exactly how such an analysis differs from one involving an 'end-determining system'.

The principal contention of the book, together with these subsidiary arguments, has unhappy consequences. The selection of 'motivational activities'—a class of behaviour including, but not co-extensive with that of consummatory response—as the definition of the subject-matter leads inevitably to a discussion of behaviour in general. For most behaviour is motivated, and a

discussion based upon consummatory activities alone would fail to raise many of the interesting problems which are arising in this field of study. Thus, one is at first surprised to find a long section on discrimination learning and generalization which has no more to do with motivation than with any other topic in the analysis of behaviour. However, since it is Dr Bindra's contention that the psychological study of motivation consists in the analysis of such activities as eating, hoarding, exploring, playing, etc., in terms of the principles of instrumental learning, such apparent digressions cannot be avoided. The book, naturally, becomes an exposition of these principles with special reference to certain types of response. A second consequence is that the motivation of behaviour, as opposed to certain instances of motivated behaviour, receives only a cursory over-confident discussion in physiological terms. The link between observations of behaviour and physiological mechanisms remains very general, vague and uncertain, despite frequent references to phase-sequences, activating mechanisms and the like. Hebb could hardly have intended his work to be more than a suggestion about how the nervous system might work in certain circumstances. Dr Bindra seems almost to accept these suggestions as non-controversial, commonplace and proven. Even if this were so, it remains unclear how, in any detail and on a particular occasion, a particular sequence of responses is produced, what brings it to an end, and how the rewarded exercise of a sequence affects its probability of recurrence. A third consequence of the principal theme is that it leads to a form of discussion in which, for instance, different aspects of eating are discussed repeatedly in different chapters. Thus, the development, the elicitation and the physiological basis for such activities are examined in widely separated places. The number of different types of behaviour involved is quite large, and, in spite of a good index and bibliography, this separation proves to be a confusing method of exposition.

The central theme of the book, that all 'motivational activities' are learned, receives considerable modification in the course of the argument. This is at first set out as a general hypothesis, and experiments such as Tolman's are quoted to suggest that the claims of ethologists and others for entirely innate responses to specific stimuli are invalid. Yet these experiments involved different conditions and species from those of the experiments quoted. It is not clear why the one should invalidate the other. It is said that the earlier experiments, showing such 'instinctive' responses, were inadequately controlled: but the nature of the inadequacy is not stated. Later in the light of other evidence concerning activities such as hoarding and with various species differences Dr Bindra concedes that there must be innate differences of 'potentiality' to make certain responses to certain stimuli. In connexion with exploratory behaviour and fear responses the claim is weakened further and it is said only that some non-specific previous learning is required. Thus, for instance, an animal will not be frightened by a particular novel stimulus until a familiar background environment has been established. And this requires learning, but not learning specific to the event in question. Finally, the claim for the acquired basis of all behaviour is restricted to mammals. It is difficult to see, at the end of these chapters, that much remains of a hypothesis that was, *prima facie*, excitingly outrageous.

There are minor defects in the book. The theoretical framework serves sometimes to give the evidence a too one-sided appearance. For instance, in the discussion of the effects of hunger on exploratory behaviour, it would be possible to cap each experiment quoted by another showing the contrary result, and, in one case, by another part of the same experiment. There are some very minor inaccuracies in the reports of the details of experiments. Expressions such as 'meaningless' are used rather carelessly of statements which are perfectly intelligible; and, to this extent, the exposition is unnecessarily confused.

It is difficult to see that the author has succeeded in his main objective—to give an account of 'motivated activities' in terms of instrumental learning and some degree of explanation in terms of physiology. The weight of evidence seems against the former, and is insufficient to sustain the latter. But it should be remembered that these are very difficult tasks and, in spite of this failure and the defects of exposition, Dr Bindra's book is a serious and informed attempt to develop a unified approach to this subject, and, as such, deserves attention.

A. J. WATSON

Behaviour and Evolution. Ed. by ANNE ROE and GEORGE GAYLORD SIMPSON. Yale University Press; London: Oxford University Press. 1958. Pp. viii + 557. 80s.

This book results from two conferences arranged jointly by the American Psychological Association and the Society for the Study of Evolution. In its twenty-three separate contributions the value and basis of the comparative approach to behaviour are discussed in a fashion always bound to the concepts of evolution. Behaviour is treated as a system of adaptation comparable with physiology or morphology.

Simpson opens Part 1 with a clear introduction to modern evolutionary theory, and then Colbert and Romer cover the course of evolution, especially that of the vertebrates, as revealed by palaeontology. Behavioural contributions here are only slight, but this part provides a phylogenetic framework for the rest.

Part 2 shows the value of comparison in illuminating the physiology of behaviour. There is a review of endocrine control by Beach and one of behaviour genetics by Caspari, which is the first to include an adequate genetical background. Sperry deals with some of the features of the developing nervous system and in particular with the extraordinary specificity of neurone connexions. Pribram and Bullock contribute on comparative neurology and neurophysiology respectively. Bullock examines the neuronal properties of various groups of invertebrates and shows that complex phenomena are found even in the simple nerve net of coelenterates. Most properties of mammalian neurones are already possessed by those of, for instance, arthropods, but Bullock believes that the real advances of the mammalian brain are based on new emergent properties of neurone groups which are little known as yet.

The broadest possible comparisons are also made by Harlow discussing the evolution of learning in Part 3. Within the mammals he finds no evidence of a clear qualitative break between man and the rest, but prefers to regard the progression as one of capacity. The whole essence of an evolutionary study is comparison, and Part 3 consists of a series of essays dealing with categories of behaviour. As a prerequisite Nissen begins by considering how we are to classify behaviour. He selects six axes of comparison (sensory capacity, locomotion, manipulation, perception, sensori-motor connexions and reasoning) and maintains that phylogeny reveals an over-all advance along them. However, at present much more is to be gained by comparing closely related species where the homologies of detailed behaviour patterns are clear (as discussed here by Hinde and Tinbergen) than by attempting to compare diverse animal groups for broader categories such as 'food-getting behaviour'. Emerson's account of the social insects is useful, but some of the other essays here smack of comparison for comparison's sake.

In Part 4, Mayr and Spieth add weight to the 'ethological' approach by demonstrating the value of behaviour as a taxonomic character and as an important isolating mechanism. Pittendrigh considers in some detail the concept of adaptation which is basic to the whole evolutionary approach and illustrates this from his own work on micro-geographical isolation in *Drosophila*.

The final section contains four attempts to apply evolutionary concepts to human behaviour and society. Washburn and Avis deal with the biological aspects of man's emergence from the primates, while Freedman and Roe consider the history of man's main psychological 'axes' (sexuality, aggressivity and acquisitiveness—a classification they find useful and capable of evolutionary analysis). Huxley writes persuasively of culture as adaptation and does not hesitate to talk of progress, and Mead discusses the rapid evolution that can occur within cultures today.

Simpson ends with a general synthesis, cutting across the parts into which the book is divided. This is very successful, and gives a measure of how well, on the whole, the individual contributions complement each other.

This book covers an immense range of topics, but never merely superficially. It will provide a useful biological background to more strictly psychological comparative psychology.

AUBREY MANNING

Contributions to Modern Psychology: Selected Readings in General Psychology. Ed. by D. E. DULANY, R. L. DEVALOIS, D. C. BEARDSLEE and M. R. WINTERBOTTOM. Oxford University Press. 1959. Pp. x + 398. 26s.

Readings in General Psychology. Ed. by PAUL HALMOS and ALAN ILIFFE. London: Routledge and Kegan Paul. 1959. Pp. ix + 251. 25s.

These two volumes arise from a problem which is painfully familiar to teachers of large introductory psychology classes: what should the students be asked to read? Clearly, one of the current introductory texts is not enough. These texts typically touch on everything and dwell on very little, often lead to students merely acquiring a repertoire of superficial jargon, and, at least in the reviewer's experience, often induce negative attitudes to psychology because students feel they are being talked down to and are ashamed to show their text, replete with photographs and cartoons, to colleagues studying 'tougher' subjects.

The editors of *Contributions* formulate an important part of this problem when they say that 'students miss... a close familiarity with the experiments and observations upon which the general discussions in textbooks and lectures are based'. They rightly observe that this sub-problem cannot be solved by sending large numbers of students in search of scattered journal articles and they produce, as one attempted solution, the present collection of supplementary readings. They have done their job well. The book contains forty articles selected from all areas likely to be taught in an introductory course; the articles have the merit of being relevant to one or other basic concept of introductory psychology, being readable, and being of immediate interest to students (who would not have to be coaxed to read most of the articles); the volume is well produced and is pleasant to handle. Each article is given a brief introduction, its published source is detailed and, at the end of the book, it is related to the relevant chapters of six current textbooks. Leaving aside the inevitable point that different editors would have produced different (although probably not better) selections of articles, the important question is: will these supplementary readings be useful to introductory courses as taught in this country? Had the book been half its price, this reviewer would have answered in an unhesitating affirmative. As it is, many teachers may feel that the monetary cost of these readings outweighs their value as a contribution to the student's studies. Suffice it to say that this volume is well worth having on the library shelf and that its possible use for introductory students merits serious consideration.

The editors of *Readings* start with the same problem of what introductory students should read but their attempted solution, for which the reviewer has great sympathy, is more radical. A 'broad-shallow' approach should be replaced by a 'narrower-deeper' one which introduces students more thoroughly to fewer topics. To implement this proposal, their book presents fifteen articles, all by British psychologists, all but one 'written within the past five years or so' and each selected to 'deal comprehensively and competently with one major subject or issue in psychology': almost exactly one-quarter of the textual material is occupied by four specially written articles. Presumably this book is to be used by assigning its articles to the student and elaborating them by means of lecture, discussion, laboratory work and perhaps, as a turn of the tables, a 'broad-shallow' text as supplementary reading. Presumably also the editors have conducted interesting teaching experiments along such lines—if so, it is a pity they make no mention of their experiences. Whether the book could be successfully used in this way is a matter on which this reviewer has grave doubts—even admitting that this matter can hardly be settled by any teacher who has not tried the book out with his own students and admitting also that teaching requirements and talents differ from one department to another.

The two solutions proposed by the editors of the present volumes deserve the attention of all psychology teachers. This reviewer would add as a third proposed solution that teachers look to paper-backs for a series of inexpensive and easily available monographs each probing at some length but not too complexly into 'one major subject or issue'. Why might it not, for example, come to pass that, in a few years time, the Penguin Psychology Series will contain a number of monographs which, collectively, provide all the reading necessary for a thoroughgoing introductory course and, subsidized as these volumes are by a wide general readership, do so at a total cost to the student of less than a pound?

I. M. L. HUNTER

Psychology of the Child. By ROBERT I. WATSON. New York: Wiley; London: Chapman and Hall. 1959. Pp. ix + 662. 56s.

The Changing American Parent. By DANIEL R. MILLER and GUY E. SWANSON. New York: Wiley; London: Chapman and Hall. 1958. Pp. xiv + 302. 52s.

Psychology of the Child is subtitled 'Personal, Social and Disturbed Child Development' and will probably be of use to American students taking courses in Personal or Social or Disturbed Child Development. The treatment is comprehensive, thorough and dull. Each chapter closes with a summary and a list of references. The book is full of meat, but it is sure tough. The English style will deter most readers; it consists almost entirely of a series of simple, rather dogmatically expressed sentences after most of which one almost expects to find True/False in brackets. It is a pity that the presentation is so much worse than the material, but 662 pages of it is more than this reviewer can recommend.

The Changing American Parent is an interesting attempt to ascertain what differences there are in methods of bringing up children among middle-class parents in the U.S.A. Two types of social context are recognized: entrepreneurial whose feature is self-help and competition; and bureaucratic whose characteristics are essentially those of the welfare state. The study took place in the Detroit area and certain predictions were made about differences in the practices of child-rearing with regard to weaning, toilet training, freedom, punishment and the like. These predictions were to a large extent fulfilled. The study is in a peculiarly American context and the findings cannot be directly translated into British conditions, but this study does indicate a significant field of investigation and discusses methodological conditions fully. There is also a chapter on the history of American views on how to bring up children which is most interesting and entertaining in itself.

JAMES MAXWELL

Hypnosis: Fact and Fiction. By F. L. MARCUSE. Harmondsworth: Penguin. 1959. Pp. 224. 3s. 6d.

L'Hypnose: Problemes Theoriques et pratiques. By L. CHERTOK. Paris: Masson. 1959. Pp. 78. 700 francs.

Popular interest in the mystery of hypnosis continues at a high level, and the Pelican is assured of a ready sale from the book stalls. As the editor, Prof. Mace, tells us, it is another contribution towards the provision for 'every boy and girl, as part of a liberal education', of 'the essential critical equipment for distinguishing the Sense from the Nonsense in Psychology and its Uses from its Abuses.'

Its author can claim to speak with some authority. A Professor in a University Department of Psychology, he has published some seventy papers in the reputable journals of psychology, and he has specialized in the study of hypnosis for over 15 years. In a formidable list of his qualifications, the editor makes no mention of post-graduate study at the Yeovil Institute. Yet he is adept in the game of deflating those with whom he does not agree, while persuading that his own scientific standards are of the very highest. Having conducted experiments with about 1000 subjects, he is one up from the start, and he is especially skilful in displaying what the editor describes as 'one of the most reliable signs of the genuine scientist: when he does not know the answer to a question he says so straight away'. Resolutely refusing to consider the phenomena of hypnosis in the context of a general psychological theory, and vigorously dismissing as unproven the explanations of those who have tried to do so, he plays fair by the popular tradition and says nothing to diminish the mystery.

The brochure by Dr Chertok makes plain reading. Although it contains little new, it is informative without being pretentious.

D. RUSSELL DAVIS

Der Thematische Apperzeptionstest (TAT). By W. J. REVERS. Bern and Stuttgart: Huber. 1959. Pp. 212. Fr./DM. 26.

Dynamische Zusammenhänge in der Psychologie. By WOLFGANG KÖHLER. Bern and Stuttgart: Huber. 1959. Pp. 125. Fr./DM. 16.80.

These are, respectively, volumes 2 and 3 of the *Enzyklopädie der Psychologie in Einzeldarstellungen*, edited by Prof. R. Heiss, Freiburg. The aim of Revers's book is to make good the lack of an introductory text-book on the TAT in the German language. The author fulfils this purpose in a very thorough way. He discusses in detail the history of pictorial testing, the rationale of the TAT (always giving due consideration to work following upon Murray's), techniques of application and evaluation, and the as yet partly unsolved problems of reliability and validation. In a final chapter he reveals deep understanding of the psychocathartic side-effects of the TAT. The author sees the TAT in a complementary role to that of the Rorschach, in that it provides the 'flesh' to the formal-structural 'skeleton' of personality revealed by the Rorschach. He proves to be well grounded, both in the practical and in the theoretical aspects of personality diagnostics.

Köhler's book appeared in English as *Dynamics in Psychology* in 1942. The salient point in this 'discours de la méthode' is the stress on a multidisciplinary approach, which in Köhler's view is essential for the psychologist. The formulations of one of the foremost thinkers in contemporary psychology always deserve the closest attention. How far, however, they come near to 'last words' in methodology remains questionable. What Köhler has not overcome is the 'compartmental' outlook of his school, though at least Lewin and Goldstein have striven to extend its concepts into the sociological and organismic-biological fields respectively.

STEPHEN KRAUSS

Psychopharmacologia. Berlin: Springer. Vol. 1, Fasc. 1, 1959. Pp. 78. DM. 9.60.

'Recent years have witnessed an unprecedented advance in research on the action and effects of psychotropic drugs, and what formerly was just a distant goal has now evolved into a new branch of science: psychopharmacology.' So begins a statement inside the front cover of this new journal, and it fairly expresses what has been happening; because of this advance, to which psychologists have greatly contributed, drugs are among other things becoming potent tools for analysing behaviour and its relations with the nervous system. The statement goes on to say that hitherto the literature has been widely scattered and that *Psychopharmacologia* hopes to overcome this. Inevitably the 'scattering' will to some extent continue, as indeed it should, since it is one of the charms of psychopharmacology that it forms links between diverse branches of biology and medicine. But a focus for publications was of course badly needed, and the new journal will make it much easier to keep up with progress and to compare the results of studying drugs from many different points of view.

The editorial policy is valiant and eclectic—'... prompt publication of scientific contributions concerned with... the effects of drugs upon behaviour, in the broadest sense of the term.' The first number contains five original papers and two reviews, all except one in English, and the standard of competence appears generally high. The topics are well balanced between old and new drugs and between animal and human experiments and clinical studies. The emphasis is, appropriately, on psychological effects; sometimes these have been studied in relation to possible underlying modes of drug action (e.g. in the papers on psychosis-like symptoms), and it will be valuable if future issues will tell us still more about relevant findings obtained by using pharmacological, neuro-physiological and biochemical techniques. In addition to articles, short 'letters to the editors' are invited, and one hopes that these opportunities will be used; book reviews and reports of symposia and other meetings might perhaps also be worth including. The printing and layout are clear and attractive.

The managing editors are Drs E. Rothlin (Basle) and A. Wikler (Lexington, Kentucky), and there is a large, varied and truly international board. In order to speed publication, the journal will appear at irregular intervals, according to the material available. This seems a good idea, especially at the beginning. The journal deserves to attract contributions of a kind which will enable it to maintain its quality and yet to appear often.

HANNAH STEINBERG

OTHER PUBLICATIONS RECEIVED

- The Appraisal Interview: Objectives, Methods and Skills.* By N. R. F. MAIER. New York: Wiley. London: Chapman and Hall. 1958. Pp. 246. 48s.
- Measurement of Learning: a Study of Verbal Rote Learning.* By M. BJÖRKMAN. Stockholm: Almqvist and Wiksell. 1958. Pp. 114. 14 Sw.kr.
- Readings in Social Psychology* (3rd ed.). Ed. by E. E. MACCOBY, T. M. NEWCOMB and E. L. HARTLEY. London: Methuen. 1959. Pp. 674. 55s.
- What is Cybernetics?* By G. T. GUILBAUD. Trans. by VALERIE MACKAY. London: Heinemann. 1959. Pp. 126. 10s. 6d.
- The Language and Thought of the Child.* (3rd. ed.). By J. PIAGET. London: Routledge and Kegan Paul. 1959. Pp. 288. 25s. (A 42-page chapter on 'Children's verbal exchanges' has been added: otherwise, both text and index are identical with the previous edition.)
- Family and Class Dynamics in Mental Illness.* By J. K. MYERS and B. H. ROBERTS. New York: Wiley. London: Chapman and Hall. 1959. Pp. 295. 56s.
- The Archetypal World of Henry Moore.* By ERICH NEUMANN. Trans. by R. F. C. HULL. London: Routledge and Kegan Paul. 1959. Pp. 138 + 107 illustrations. 35s.
- The Society of Captives: a Study of a Maximum Security Prison.* By G. M. SYKES. Princeton University Press. London: Oxford University Press. 1958. Pp. 144. 30s.
- The Language of Psychology: a systematic, full-length treatment of scientific language, definition, and theory construction in psychology.* By G. MANDLER and W. KESSEN. New York: Wiley. London: Chapman and Hall. 1959. Pp. 301. 54s.
- A Man for Every Woman.* By R. H. KLEMER. New York: The Macmillan Company. Pp. 193. 19s.
- Psychology, the Nurse and the Patient* (3rd ed.). By D. M. ODLUM. London: Iliffe and Sons. 1959. Pp. 200. 15s.
- Sigmund Freud's Mission.* By ERICH FROMM. London: George Allen and Unwin. 1959. Pp. 120. 12s. 6d.
- The Psychology of Affiliation: Experimental Studies of the Sources of Gregariousness.* By S. SCHACHTER. Stanford University Press. London: Oxford University Press. 1959. Pp. 141. 30s.
- Metamorphosis: On the Development of Affect, Perception, Attention, and Memory.* By ERNEST G. SCHACHTEL. New York: Basic Books. 1959. Pp. 344. \$6.00.
- Group Psychoanalysis.* By B. B. WASSEL. London: Peter Owen. 1959. Pp. 306. 30s.
- A Search for Man's Sanity: the Selected Letters of Trilby Burrow with Biographical Notes.* Ed. by EDITORIAL COMMITTEE OF THE LIFWYNN FOUNDATION. Oxford University Press. 1958. Pp. 615. 60s.
- Introduction to Counseling.* By E. L. TOLBERT. New York and London: McGraw-Hill. 1959. Pp. 322. 46s.
- Aion: Researches into the Phenomenology of the Self.* Vol. 9, Pt. II, Collected works of C. G. JUNG. By C. G. JUNG. London: Routledge and Kegan Paul. 1959. Pp. 344. 32s.
- Tierpsychologische Beiträge zur Phylogense der Ich-über-ich-Instanzen.* By ADAM ZWEIG. Bern: Hans Huber. 1959. Pp. 82. D.M. 12.50.
- Exploration in Group Relations.* By E. L. TRIST and C. SOFER. Leicester University Press. 1959. Pp. 68. 7s. 6d.
- The Sixth Sense: An Inquiry into Extra-sensory Perception.* By ROSALIND HEYWOOD. London: Chatto and Windus. 1959. Pp. 224. 21s.

- Water Witching U.S.A.* By E. Z. VOGT and R. HYMAN. University of Chicago Press. London: Cambridge University Press. 1959. Pp. 248. 37s. 6d.
- Complex, Archetype, Symbol.* By J. JACOBI. London: Routledge and Kegan Paul. 1959. Pp. 230. 21s.
- The Effect of Pharmacological Agents on the Nervous System.* Ed. by F. J. BRACELAND. (Res. Publ. Ass. Nerv. Ment. Dis. Vol. XXXVII.) London: Baillière, Tindall and Cox. 1959. Pp. 488. 108s.
- The Sociological Imagination.* By C. W. MILLS. Oxford University Press. 1959. Pp. 234. 36s.
- A Changing America.* By A. W. ZELOMEK. New York: Wiley. London: Chapman and Hall. 1959. Pp. 181. 32s.
- Critical Incidents in Psychotherapy.* Ed. by S. W. STANDAL and R. J. CORSINI. New York: Prentice-Hall. 1959. Pp. 396. \$6.95.
- Men who Manage.* By M. DALTON. New York: Wiley. London: Chapman and Hall. 1959. Pp. 318. 54s.
- Molecules and Mental Health.* By F. A. GIBBS. Philadelphia: Lippincott. London: Pitman Medical Publishing. 1959. Pp. 189. 40s.
- Art and the Creative Unconscious.* By ERICH NEUMANN. London: Routledge and Kegan Paul. 1959. Pp. 232. 25s.
- Seelische Gesundheit.* By W. BETTSCHART, H. MENG and E. STERN. Bern: Hans Huber. 1959. Pp. 356. Fr./DM 34.
- Yoga.* By E. WOOD. Harmondsworth: Penguin. 1959. Pp. 272. 3s. 6d.
- Fear, Punishment and Anxiety and the Wolfenden Report.* By C. BERG. London: Allen and Unwin. 1959. Pp. 126. 20s.
- Write and Reveal: Interpretation of Handwriting.* By P. FRIEDENHAIN. London: Peter Owen. 1959. Pp. 184. 30s.
- The Archetypes and the Collective Unconscious.* By C. G. JUNG. Vol. 9, Pt. 1, of *Collected Works*. London: Routledge and Kegan Paul. 1959. Pp. 451. 52s. 6d.
- Bilingualism and Intelligence.* By W. R. JONES. University of Wales Press. 1959. Pp. 67. 4s. 6d.
- The Priest and the Sick in Mind.* By A. A. A. TERRUWE. London: Burns and Oates. 1959. Pp. 123. 16s.
- Ciba Foundation Study Group No. 1—Pain and Itch: Nervous Mechanisms.* Ed. by G. E. W. WOLSTENHOLME and M. O'CONNOR. London: Churchill. 1959. Pp. 120. 12s. 6d.
- Handicapped Children in Britain: Special Subject List No. 30.* Compiled by W. A. AXFORD. London: The Library Ass. 1959. Pp. 53. 4s.
- Logique, Apprentissage et Probabilité.* By L. APOSTEL, A. R. JONCKHEERE and B. MATALON. Paris: Presses Universitaires de France. 1959. Pp. 186. 900F.
- Apprentissage et Connaissance.* By P. GRÉCO and J. PIAGET. Paris: Presses Universitaires de France. 1959. Pp. 185. 900F.
- Les Fonctions Grammaticales Chez L'Enfant.* By J. WITTWER. Neuchâtel: Delachaux et Niestlé. 1959. Pp. 295. Fr.s. 12.
- Psychology of Adolescence.* By LUELLA COLE. 5th ed. New York: Rhinehart. 1959. Pp. 731. \$7.00.
- A Laboratory Guide in Psychology.* By B. N. MUKHERJEE. India, Patna: Pustak-Bhandar. 1956. Pp. 104. 10 Rupees.
- Le CAT (Children Apperception Test).* By P. MULLER. Berne: Huber. Paris: Presses Universitaires de France. 1958. Pp. 146. No price given.
- Die Aggressivität bei Schülern.* By R. MANDEL. Bern: Hans Huber. 1959. Pp. 115. Fr./DM. 14.

PROBLEM-SOLVING, PAST EXPERIENCE AND AVAILABILITY OF FUNCTIONS

BY PER SAUGSTAD AND KJELL RAAHEIM

Psychology Department, University of Oslo

Following on previous studies which explored problem-solving in terms of availability of functions, the present experiment considered whether the necessary past experiences for the solution of a problem can be conveyed to subjects by demonstrating to them the functions found to be critical for the solution. The 'ball problem' was given to sixty-five 17-year-old school boys. The two critical functions were demonstrated beforehand to twenty boys of whom nineteen solved the problem within the 30 min. allowed. Of the forty-five boys given no such demonstration only ten attained solution. These results support the conclusion drawn from earlier experiments that a subject will solve a problem if he has the necessary functions available. They also support the more general contention that a fundamental aspect of problem-solving is uncovered by considering the process in terms of the functions of concrete objects. The results are related to a number of experiments on the effect of hints at demonstration on problem solving.

I. INTRODUCTION

Evidently the higher intellectual processes in man are determined by an enormous amount of past experience. An adequate understanding of thinking, therefore, is inconceivable without a conceptual framework which makes it possible to handle the effect of past experience. American theorists of learning (Thorndike, 1931; Watson, 1924; Guthrie, 1935; Skinner, 1938; Spence, 1932; Hull, 1943) have attempted to provide such a framework in terms of principles of learning derived from a study of some relatively simple situations. However, it seems fair to state that this group of psychologists has neglected the problems of perception. On the other hand, Gestalt psychologists (Wertheimer, 1945; Koehler, 1917; Koffka, 1935; Lewin, 1946; Székely, 1950*a, b*, 1958) have attempted to make a framework in terms of principles of perception derived from a study of relatively simple perceptual situations. This second group apparently has neglected the problems of past experience and learning (Saugstad, 1958). Evidently thinking must be understood in terms of learning as well as of perception and also in terms of principles for memory. The one-sided approaches to the study of thinking among experimental psychologists may be due to the fact that, during the last generation, the thought processes have relatively seldom been the object of a more direct experimental analysis. Attempts are usually made to account for them in terms of principles derived from studies of simpler situations particularly within learning and perception.

It is a conviction of the writers that not only will an understanding of thinking be increased by a direct experimental attack on problems in this area, but so also will the processes involved in learning and perception. An outline for a framework based upon a more direct attack on problems of problem-solving has been presented elsewhere (Saugstad, 1955, 1958) and will be briefly indicated here.

In the hope of making communication more efficient—an overwhelmingly difficult problem in studies in this area—attention is restricted to problem situations involving concrete objects, such as stones, sticks, ropes, balls, iron nails, boxes, pieces of string

and implements of various sorts (e.g. pairs of pliers, screwdrivers, hammers). In problem situations of such a concrete nature the solution is attained by the use of definite objects in definite ways and, in the attainment of the solution, the objects may be said to serve definite functions. A concrete problem may now be subjected to a logical analysis in terms of these functions. The solution may be conceived of as the arrangement of a number of functions into some definite sequence. In previous experiments (Saugstad, 1955; Saugstad & Raaheim, 1957), the functions available in an individual with regard to a definite object have been assessed by the following simple procedure, called an availability test. The individual is presented with one object at a time and asked, according to some specified rules, to list as many functions for this object as possible. This procedure is carried through without any mention of a definite problem situation.

So far, the term 'function' has been used here in its relatively vague every-day meaning. Instead of attempting to lay down a more formal definition at present, it is thought wiser to attempt to clarify the term slowly as the experimental work progresses. To some extent its meaning is implicitly given by the procedure adopted in the first experiment by Saugstad (1955), but to avoid some misunderstandings (Székely, 1958; Saugstad, 1958) some brief comments on certain aspects will be presented.

Functions may be considered as varying with regard to specificity or generality. One may think of a function as something determined by some very specific context or one may have in mind a large variety of situations. For example, I may use a stone in one specific situation to crack one specified nut, or I may use a stone to crack nuts in general in a variety of situations. This wider generality was what Saugstad (1955) had in mind when he designed his first availability test in such a way that the subjects were asked, first, to state a function in a general form and then to give three concrete examples of the function. The procedure followed in the experiments performed thus rests on the assumption that functions may be perceived or conceived of in a more or less general form. The results of the experiments performed within this framework indicate that, under certain circumstances, functions may be regarded as being relatively independent of the situational context. On the other hand, experiments by Duncker (1945), Adamson (1952), Birch & Rabinowitz (1951) reveal that, under other circumstances, availability of a function may be highly dependent upon the situational context. The degree to which a function is dependent upon the situational context can be clarified only by more systematic empirical research. Attention should be called to the fact that, so far, attempts have been made within this framework to stay away from situations which may be conducive of fixation of a function.

In connexion with the problem of specificity or generality of functions, it should be pointed out that the present approach grew out of a dissatisfaction with the currently dominant approaches to learning and problem-solving. On the one hand, there is the approach of Gestalt-psychology where the term 'cognitive or thought structure' so strongly stresses the situational context that the situation can hardly be analysed into subunits. Evidently human beings and primates generally are, to a considerable extent, able to isolate and manipulate definite aspects of a stimulus situation independent of the context. In this connexion, it should be emphasized that, to a large extent, we conceive of concrete objects as discrete. On the other hand, the various approaches

in terms of 'stimulus-response' tend to neglect the situational context. This makes it difficult, if not impossible, to predict over a variety of situations what is to be considered as the stimulus. It is hoped that the term 'function' might strike a more happy medium than either the term 'cognitive structure' or the term 'stimulus' as used by a number of behaviourists. It seems considerably easier to isolate and manipulate the referents of the term 'function' than those of the term 'cognitive structure'. Also, in contrast to the term 'stimulus' which tends to be conceived in terms of physical energy and in some cases in terms of physiological processes, the term 'function' will probably cover a larger variety of situations since it has a more definite phenomenological anchoring.

A number of hypotheses with interesting theoretical implications may be formulated with regard to functions. One might, for example, formulate hypotheses relevant to the problem of conditions affecting availability of functions, or hypotheses relevant to the problem of how individuals proceed to arrange a number of functions into the correct sequence for the solution of a problem, or hypotheses concerning the relationship between the ability to produce functions and other types of intellectual abilities. In the experiments performed so far, concentration has been on the following question: will a person who has available in a problem situation the necessary functions for the solution of the problem actually solve this problem? In the exploration of this question the assumption has been made that a function available in the availability test will be available also in the problem situation. Results of the experiments (Saugstad, 1955, 1957; Saugstad & Raaheim, 1957) suggest that the question is to be answered in the affirmative. An individual having available the necessary functions will solve the problem. Some of the theoretical implications of this result have been discussed elsewhere (Saugstad, 1958). Here, it should just be pointed out that it is reasonable to state that an individual, who has available the necessary functions for the solution of a problem, is in possession of the necessary past experience for solving the problem. If experience is determined in this way, the results may be stated as follows: an individual having available the necessary past experiences for the solution of a problem will solve this problem.

The problem of whether or not availability of the necessary past experiences is sufficient for solution is of considerable theoretical interest. Mainly based upon Maier's (1930) well-known pendulum problem, Gestalt psychologists (Wertheimer, 1945; Koffka, 1935) and Gestalt-oriented psychologists (Leeper, 1951) have argued that availability of the necessary past experiences is not sufficient for the solution of a problem. In his experiment, Maier demonstrated to a group of subjects what he considered to be the three essential principles for the solution of the pendulum problem. Since only one in a group of subjects after this demonstration solved the pendulum problem, Maier concluded that availability of the necessary past experiences for the solution of a problem was not sufficient for arriving at the solution. A repetition of this experiment by Saugstad (1957) took steps to make sure that subjects had understood the demonstrations, as purported by Maier and as necessitated by his conclusion. It revealed that close to 100% of the subjects having the principles available solved the problem. The principles involved in Maier's demonstrations were very abstract and this may be the reason why the demonstrations had little or no effect. It might be thought that a demonstration, by way of concrete objects, of the

necessary functions for solution might be more effective in conveying to the subjects the necessary past experiences. If the subject having received such a demonstration solved the problem, this would be an effective way of showing that an analysis of problem-solving in terms of functions reveals an important aspect of problem-solving behaviour. Such a type of experiment will also have some bearing on the problem of the effects of hints and demonstrations in problem-solving which has been the subject of a number of investigations (Maier, 1930, 1931, 1933, 1945; Reid, 1951; Battersby, Teuber & Bender, 1953; Marks, 1951; Cofer, 1951).

II. EXPERIMENTAL

(i) *Subjects*

All the pupils in one class in a secondary school in Oslo served as subjects. The pupils, males aged 17-18 years, were majoring in mathematics and physics. The experiment was planned with two classes, but because the results seemed to be highly reproducible, it was discontinued when one class had been tested.

(ii) *The problem*

Since the problem has been described in detail by Saugstad & Raaheim (1957), only a rough outline will be given here. The subject is faced with the task of transferring a number of steel balls from a glass into a metal cylinder. The glass containing the steel balls is placed on a movable wooden frame and the metal cylinder to the side of it. The subject has to keep behind a wooden table, or a chalk line, 260 cm. away from the wooden frame. At his disposal for the solution of the problem the subject has four newspapers of a definite size, a length of string, a pair of pliers, five elastic bands, and one iron nail. The correct solution to the problem consists in bending the nail by the pair of pliers into a hook, attaching the nail to the string and throwing the nail into the wooden frame, making it catch onto the frame, and then pulling in the frame on which the glass of steel balls lies. The newspapers are then to be rolled into tubes, which are inserted into each other to form a continuous passage. The newspapers are held together by the elastic bands. Through this tube the steel balls are directed into the cylinder from behind the table.

(iii) *Procedure*

Demonstration of the critical functions. Analysis of the problem had previously shown that all subjects giving the function 'hook' for the nail and the function 'tube' for the newspaper, would solve the problem. It was decided to demonstrate only these two functions, called the critical functions. The demonstrations were carried out by reading the following instructions. 'We will now present you with some objects which all may be used in a number of ways. We want you to give us some examples for the use of the objects.' A newspaper rolled into a tube is handed to the subject. 'This object you may use to conduct something through. Could you give some examples?' A nail bent into a hook is handed to the subject. 'This object may be used to catch things with. Could you give some examples?' When the subject had given as an example in answer to the first question: air, or some other gas, water,

sand, balls, pebbles, or some other smaller objects, this was considered a satisfactory indication that he had grasped the meaning of the function stated for the newspaper. When the subject had answered the second question with: angle for fishing, to hang things on, as a hook, or to catch on to wood or some softer material, the example was considered an indication that he had grasped the meaning of the function stated for the bent nail. If the subject did not give any of the examples mentioned he was asked to continue until he had met the criterion as stated.

Design. The subjects were first given the demonstrations as described in one room and then immediately led into another room where they were presented with the problem situation. The control group of an experiment previously reported by Saugstad & Raaheim (1957) served as control group for the present study. This group was presented with the problem without any previous demonstrations. The subjects of this group, forty-five pupils from secondary schools in Oslo, had the same characteristics as the subjects in the experimental group. The subjects of both groups were tested individually.

Subjects were allowed 30 min. on the problem. The demonstrations of the critical functions took from 1 to 3 min.

III. RESULTS

Table 1 shows the number of subjects, in each group, who had solved the problem after 5, 10, 15, 20, 25, or 30 min. (The percentage number of solutions is shown in brackets, to the nearest whole percent.) The figures show that, at every time interval, the subjects who had received the demonstrations were superior to those who had not. Taking 30 min. as the time limit, solutions were given by nineteen out of the twenty subjects of the experimental group as opposed to only ten out of the forty-five subjects of the control group: this difference is significant at the 1 % level (χ^2 corrected for continuity).

Table 1

Group	No. of solutions						Total no. of subjects
	5 min.	10 min.	15 min.	20 min.	25 min.	30 min.	
Exp.	5 (25)	14 (70)	17 (85)	18 (90)	19 (95)	19 (95)	20
Contr.	0 (0)	0 (0)	3 (7)	7 (16)	9 (20)	10 (22)	45

Only two subjects in the experimental group spent more than 20 min. on the solution. As noted above the demonstrations took from 1 to 3 min. Therefore, the difference in number of solutions between the experimental and the control group cannot be explained in terms of greater familiarity with the material on the part of the subjects in the experimental group.

The results which seem to be highly reproducible, indicate that subjects from the same population as the subjects of this experiment will almost always solve the ball problem when the two critical functions have been demonstrated to them.

IV. DISCUSSION

Almost all the subjects who had the two critical functions demonstrated to them solved the problem. This suggests that an analysis in terms of availability of functions uncovers an important aspect of problem-solving behaviour. By means of this analysis, it appears to be possible, in a simple way, to provide an individual with the part experiences necessary for the solution of a problem. The analysis of a problem situation in terms of functions makes it possible to predict which aspects in a problem situation should be pointed out to the subjects not solving the problem. The present results thus make it possible to formulate, in a tentative manner, the conditions under which a hint will be effective: a hint will be effective to the extent to which it makes available the critical functions for the solution of a problem.

We will from this angle consider some experiments with concrete objects where hints are introduced. Maier (1931) constructed the problem of tying together two ropes which were hung so far apart that, when an individual was holding one, the other could not be reached. To solve this problem, the subject had to make a pendulum out of one of the ropes by tying an object onto it. In this problem situation, it was found effective to give one of the ropes a push. The push may reasonably be said to make available the function served by the rope. Battersby *et al.* (1953) used the same problem situation, but they gave hints by placing objects which suggested the function of a weight to the subject—thus emphasizing another aspect for the construction of the pendulum. Again it is to be expected, as it transpired, that the hint was effective. Cofer (1951) also used the same situation but had his subjects memorize lists of words before they were presented with the problem. More solutions were produced in a group of subjects who had memorized among other words, rope, swing, and pendulum.

The above formulation for the effect of hints may also help to clarify some of Maier's (1945) reasoning around a more recent experiment. We will concentrate on one aspect of this experiment. Maier presented one group with the problem of constructing a hat rack. This was done by making a longer pole out of two shorter ones, and by placing this longer pole between floor and ceiling and clamping them together by a table clamp. The clamp in this construction served as a hat rack. Another group of subjects were placed before the above-mentioned construction after this had been completed and asked to list possible functions for this construction. It was found that there was a proportionally larger number of subjects in the first group who were able to construct the hat rack than there were subjects in the last group who could suggest the function of a hat rack or some similar function. As we understand him, Maier argues that the situations for the two groups were equivalent and that it is strange, therefore, that the results came out as they did. But the two situations are not equivalent and the results are not at all surprising. Evidently, it is difficult to perceive or conceive of the function of a hat rack in the clamp; few subjects in the second group perceived this function. The first group being assigned the task of constructing a hat rack has, *through the instructions*, received a hint that some object may serve as a hat rack. For this reason, a larger proportion of them finds the function of hat rack for the clamp. It is thus seen that the two groups are treated entirely differently in one crucial aspect. The situation for the subjects receiving the

instruction to build a hat rack is similar to that of the subjects in Cofer's (1951) experiment who, before they were presented with the problem of tying the two ropes together, had memorized the list containing the critical words, rope, swing, pendulum. In both experiments, the subjects received important verbal hints concerning the critical function for one crucial object in the attainment of the solution to the problem.

In still another of his ingenious experiments Maier (1933) was able to raise the number of solutions in a group of subjects by giving them a lecture on how to solve problems. The results of this experiment are of course very interesting, even if it is difficult to draw any definite conclusions as to what in the lecture had the effect of helping the subjects to find more solutions. According to the present analysis in terms of functions, it is natural to assume that the lecture helped the subjects find more functions than they otherwise would have found. In view of the content of the lecture this interpretation seems reasonable.

The present results also serve as a check on one of the assumptions underlying the conclusion drawn from previous experiments by Saugstad (1955) and Saugstad & Raaheim (1957) that a subject having available the necessary functions in the problem situation will solve the problem. Some of the main assumptions underlying this conclusion have recently been discussed elsewhere by Saugstad & Raaheim (1959). One of these had a bearing on the fact that the availability tests might have a selective effect. The subjects who, according to the availability tests, were credited with the necessary functions might be superior in general intelligence or in some problem-solving ability. Since they were superior in one or both of these respects they might be superior in combining the necessary functions into the correct sequence involving the solution. Thus, there might possibly exist a group of subjects having the necessary functions available in the problem situation, but not being able to combine the functions into the correct sequence. Previous analysis (Saugstad & Raaheim, 1957) makes it probable that there exists *no* such group of subjects. The results of the present experiment give a further check on this assumption. In the experimental group, close to 100 % solved the problem as against 20 % in the control group. The problem solvers in the first group can, therefore, not reasonably be considered superior to the problem solvers of the second group with regard to a possible problem-solving ability or to a possible general intelligence. When almost all subjects of the experimental group solved the problem the reason is most likely that the main determiner of success in problem-solving is availability of functions. However, it should be noted that the combination of the functions for the subjects in the experimental group might be somewhat easier, since the previous demonstrations might have emphasized the two critical functions and thus made easier the combining of them into the solution. Yet it seems safe to conclude that there were still a number of possible functions which might be combined. If the ability to combine functions or some sort of parts was an important factor in determining success in problem-solving one would expect it to manifest itself also in the problem situation of this experiment.

The writers are indebted to the Norwegian Research Council for Science and the Humanities for a grant to carry out these Investigations, to cand. psychol. Johannes Hovden for help in collecting the data, and to Prof. William W. Lambert, Cornell University, for criticisms of the manuscript.

REFERENCES

- ADAMSON, R. E. (1952). Functional fixedness as related to problem solving: a repetition of three experiments. *J. exp. Psychol.* **43**, 288-91.
- BATTERSBY, W. S., TEUBER, H. L. & BENDER, M. B. (1953). Problem-solving behavior in men with frontal or occipital brain injuries. *J. Psychol.* **35**, 329-51.
- BIRCH, H. G. & RABINOWITZ, H. S. (1951). The negative effect of previous experience on productive thinking. *J. Exp. Psychol.* **41**, 121-5.
- COFER, C. N. (1951). Verbal behavior in relation to reasoning and values. In *Groups, Leadership and Men*. Ed. by H. Guetzkow. Pittsburgh: Carnegie Press. (Quoted after Johnson, 1955.)
- DUNCKER, K. (1945). On problem-solving. *Psychol. Monogr.* **53**, no. 5 (whole no. 270).
- GUTHRIE, E. R. (1935). *The Psychology of Learning*. New York: Harper.
- HULL, C. L. (1943). *Principles of Behavior*. New York: Appleton.
- JOHNSON, D. M. (1955). *The Psychology of Thought and Judgment*. New York: Harper.
- KOEHLER, W. (1917). *Intelligenzprüfungen an Anthropoiden*. I. Berlin: König. Akad. Wiss.
- KOFFKA, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.
- LEEPER, R. (1951). Cognitive processes. In *Handbook of Experimental Psychology*. Ed. by S. S. Stevens. New York: Wiley.
- LEWIN, K. (1946). Behavior and development as a function of the total situation. In *Manual of Child Psychology*. Ed. by L. Carmichael. New York: Wiley.
- MAIER, N. R. F. (1930). Reasoning in humans: I. On direction. *J. Comp. Psychol.* **10**, 115-43.
- MAIER, N. R. F. (1931). Reasoning in humans. II. The solution of a problem and its appearance in consciousness. *J. Comp. Psychol.* **12**, 184-94.
- MAIER, N. R. F. (1933). An aspect of human reasoning. *Brit. J. Psychol.* **24**, 144-55.
- MAIER, N. R. F. (1945). Reasoning in humans. III. The mechanisms of equivalent stimuli and of reasoning. *J. Exp. Psychol.* **35**, 349-60.
- MARKS, M. R. (1951). Problem solving as a function of the situation. *J. Exp. Psychol.* **41**, 74-80.
- REID, J. W. (1951). An experimental study of 'analysis of the goal' in problem-solving. *J. Gen. Psychol.* **44**, 51-69.
- SAUGSTAD, P. (1955). Problem-solving as dependent on availability of functions. *Brit. J. Psychol.* **46**, 191-8.
- SAUGSTAD, P. (1957). An analysis of Maier's pendulum problem. *J. Exp. Psychol.* **54**, 168-79.
- SAUGSTAD, P. (1958). Problem-solving and availability of functions. A discussion of some theoretical aspects. *Acta Psychol.* **13**, 384-400. *Nord. Psychol.* **10**, 216-32.
- SAUGSTAD, P. & RAAHEIM, K. (1957). Problem-solving and availability of functions. *Acta Psychol.* **13**, 263-78. *Nord. Psychol.* **9**, 205-20.
- SAUGSTAD, P. & RAAHEIM, K. (1959). Problem-solving and availability of functions in children. *Acta Psychol.* **16**, 45-58. *Nord. Psychol.* **11**, 45-58.
- SKINNER, B. F. (1938). *The Behavior of Organisms*. New York: Appleton.
- SPENCE, K. W. (1932). Gradual versus sudden solution of discrimination problems by chimpanzees. *J. Comp. Psychol.* **25**, 213-24.
- SZÉKELY, L. (1950a). Knowledge and thinking. *Acta Psychol.* **7**, 1-24.
- SZÉKELY, L. (1950b). Productive processes in learning and thinking. *Acta Psychol.* **7**, 388-407.
- SZÉKELY, L. (1958). Some comments on problem-solving availability and test-magic. *Acta Psychol.* **14**, 152-7. *Nord. Psychol.* **10**, 108-13.
- THORNDIKE, E. L. (1931). *Human Learning*. New York: Century.
- WATSON, J. B. (1924). *Psychology from the Standpoint of a Behaviorist* (2nd. ed.). Philadelphia: Lippincott.
- WERTHEIMER, M. (1945). *Productive Thinking*. New York: Harper.

(Manuscript received 3 February 1959)

REGRESSION TO THE MEAN—A CONFUSED CONCEPT

BY A. D. B. CLARKE, ANN M. CLARKE AND R. I. BROWN

Psychology Department, The Manor Hospital, Epsom, Surrey

It is suggested that psychologists, like other people, are not entirely immune from the tendency to 'blind themselves with science'. Statistics in particular provide a convenient *mystique* and thus an encouragement to use catch-phrases in describing events we do not fully understand. One typical example is the phrase 'regression to the mean', often used to explain changes in extreme test score and usually regarded as a wholly statistical problem without psychological implications. This paper, however, by reference to physical and psychological data, attempts to indicate its primarily behavioural bases and to show why groups chosen for study on the basis of extreme test score will tend to regress towards the mean of the general population, while regression of extreme scores in groups not so selected will be towards the mean of their own specific population. There appear to be two main factors at work. (i) The effects of fleeting, inconstant personal alterations in response to the test (the so-called 'errors of measurement' which are more complex than they sound). (2) The essentially non-linear nature of most individual physical or psychological growth. To say without further qualification that test scores have changed 'because of regression' gives no greater information than to state that they have changed because they have changed.

I. STATEMENT OF THE PROBLEM

The concept of regression towards the mean,* which in psychology is intimately linked with the problem of errors of measurement, arises from the following established facts. (1) The correlation between test and immediate retest, which theoretically should be perfect, in practice never reaches unity. (2) The tendency of extreme scores at both ends of the normal distribution to move towards the mean of the distribution on retest. (3) The fact that test-retest correlations tend progressively to diminish with time. Below the age of 20, this diminution is in the order of 0.04 per year, excluding children under 5 years of age, where it is larger (see Thorndike, 1933, 1940; Eysenck, 1953).

It is the purpose of this article to present in some detail factors to account for these observed findings.

II. INTRODUCTION

One has only to read *Psychological Abstracts* to become painfully aware that, like others, we psychologists are not entirely immune from the tendency to blind ourselves with science. Statistics in particular provide a convenient *mystique* and thus an encouragement to use catch-phrases in describing events we do not fully understand. One typical example is the phrase 'regression to the mean': a tendency for extreme scores to move towards the average on retest is often tersely described and explained in these terms, with the implication that such changes are statistical, involve only errors of measurement, are of no psychological significance and merit little further discussion. It is of course true that it cannot be considered apart from statistics, but after all, statistics represent measures of behaviour, and it is to behaviour primarily

* No attempt is made here to deal with "biological regression", that is, regressive changes between generations; in this paper we are only concerned with intra-individual or intra-group changes. Nor do we aim to discuss regression coefficients and the prediction of a given variable value from its known relationship with another.

that we must look for explanations of regression. In this paper, therefore, we attempt to comment on the concept, to show that considerable confusion has existed and to indicate its behavioural bases. Our suggestions, it must be admitted, have emerged only slowly and painfully from the background of our own—and we suspect other people's—confusion. It is not suggested, however, that the principles we have advanced are novel, although we have not seen them presented in detail elsewhere, and it seems likely that they will be unfamiliar to many psychologists.

III. PRESENT USE OF THE CONCEPT

The following are examples taken from the literature.

(1) McNemar (1940) cites 54 children, between I.Q. 140 and 149, who on retest a week later lost an average of 5 points which he attributed to regression resulting from 'errors of measurement' from Binet, Form L to Binet, Form M.

(2) The same writer (McNemar, 1940) indicates that, 'unless variability increases from test 1 to test 2, the correlation between initial I.Q. and gains *must* be negative since in practice r_{12} will never reach unity'. This seems to be a sophisticated way of saying that unless people move away from the mean on test 2, they will move towards it. He adds, however, that 'this of course does not explain the negative correlation. . . .'

(3) Terman's group of gifted children, chosen on the basis of an initial I.Q. of 140 or above showed an average decrease on later testing due, in McNemar's opinion, solely to regression.

(4) Eysenck (1953) shares this view about the Terman study, stating that: 'as a whole the group was still highly superior intellectually, the boys having dropped slightly in I.Q. and the girls having dropped somewhat more, a regression effect towards the mean which might have been expected on purely statistical grounds. . . .'. The literal-minded might perhaps be forgiven if he interpreted this as indicating that pure statistics might possess an anti-feminist bias.

(5) Discussing this same problem, Miles (in Carmichael, 1954) states that detailed analysis of various factors failed to reveal the causes of the score decrement. The score decrease in the boys was attributed to 'a difference in the tests used or to simple regression; for the girls the larger decrement was attributed to these factors or to a change in developmental rate and age of intellectual maturation. A sex difference in social motivation might also have been involved'. A similar differential decrement occurred on achievement test quotients. . . . 'Here again test standardization and regression factors were probably involved.'

(6) The same writer (Miles, 1954) discusses various follow-up reports concerning late adolescent or young adult mental status of gifted children after 6–16 years from their first selection. Over 80 % maintained their position while all the remainder were in high centiles, and 'none regressed to the mean'.

(7) Cronbach (1949) states that: 'A person's earned score on a test is his true score, plus or minus some chance error. The *true score* is the average we would get if we tested him over and over, ironing out the unpredictable variations. Of course "good luck" makes the raw score higher than the true score, and "bad luck" makes the raw score lower. Then if we find a person with a very high score, more likely than not, part of that score is due to good luck. If we test him again, we can expect him

to slip towards his true score since chance rarely favours the same person continually. The general rule for regression is: scores above average will more often decline than increase on retesting; scores below average will tend to increase on retesting. When a person is found to have an I.Q. of 150, then his 'true' I.Q. is likely to be a few points lower. *How far scores shift on retesting depends on the reliability coefficient.* [Our italics.]

(8) Vernon (1951) commenting on Schmidt's (1946) rightly criticized study of I.Q. gains in defective children, states that: 'The correlation between tests applied to a representative group four and a half to seven and a half years apart would hardly exceed 0.60, and might be lower. *Hence, due to regression effects alone, we would expect an apparent rise in average I.Q. of 19 points, that is from 52 to 71. The statistically untrained person finds this very difficult to grasp, but it is really quite simple. If a representative group was tested (with accurately standardized tests) at 12 and 16 years, there would be as many children with I.Q.'s below 70 on the second as on the first occasion. But they would not all be the same children. Owing to imperfect correlation some of those scoring above 70 at 12 fall below 70 at 16, and vice versa. If, therefore, we pick out only those below 70 at 12 years [as Miss Schmidt has done] they inevitably show an average rise by 16; this rise has no psychological significance at all since we are neglecting the above 70's who show a fall in the same period. . . .*' [Our italics.]

(9) Clarke, Clarke & Reiman (1958) state that 'Statistical regression towards the mean is another concept frequently used to account for I.Q. changes; where this concerns errors of measurement on the same or different tests, or where different tests have a low initial correlation when given within a day or two of each other, then this is entirely correct. In such cases, regression of extreme scores is inevitable and of no psychological significance'. As will be seen later, errors of measurement are primarily behavioural and therefore it is incorrect to state (as in this and example 8) that they are of no psychological significance.

These quotations give one little help in understanding what regression really is; instead, if some of the statements are taken literally (particularly those emphasized by our italics), the impression may be gained that often the statistical cart has been put before the behavioural horse, and that in the test situation the person and his score are puppets manipulated by the strings of r . A similar point has been made by Burt in an unpublished Note on the Concept of Regression which he kindly sent us. He writes that some authors discuss regression 'as though regression itself was a physical cause'.

In summary, then, the following points emerge from the literature.

(a) Changes in test score 'due to regression' are the result of errors of measurement (usually undefined) and are psychologically meaningless.

(b) Such changes are thought of as being primarily of a statistical nature.

(c) The greater the distance from the universe mean, the greater will be the regression.

(d) Regression changes can be exceedingly large: see Vernon's (1951) view that 19 points of the increments claimed by Schmidt could be accounted for by regression.

(e) Time between test and retest seems to be considered relevant to the amount of regression; thus in example 8 we find 'The correlation . . . over 4 and a half years . . . would hardly exceed 0.60 . . . Hence due to regression effects alone . . .'

The problem of regression to the mean is one aspect of the larger problem of test reliability, and it is appropriate, therefore, to recapitulate briefly some points made elsewhere (Clarke, *et al.* 1958). Two quite distinct aspects are involved: (1) the tendency for the test to measure the same quality in the same way on different occasions; and (2) the tendency for the quality measured to remain the same at different points in time. The first is best measured by immediate or almost immediate retest,* which will tend on the average to give a maximal estimate of 'errors of measurement', and the second by retesting at varying time intervals. In general it is found that the longer the time interval the greater the tendency to individual change and hence the lower the correlation coefficient. Eysenck (1953), for example, calculated from Thorndike's (1933, 1940) data that the fall-off in correlation tended to be linear, averaging 0.04 per year; since there is no reason to suppose that errors of measurement are cumulative, this must be due to other factors.

The only article known to the writers which approaches a satisfactory description of regression is that by McNemar (1940), although the explanatory bases remain unstated. He gives examples from the fields of mental and physical measurement and indicates for instance that, if a number of 8-year-old children are selected for unusual height and are then measured again at the age of 16, their mean height will then be nearer the universe mean than 8 years earlier. If, however, a number of children of Swedish extraction were similarly measured and re-measured, at both ages their mean height will be half a standard deviation above the universe mean (presumably for genetic reasons) and regression of extreme scores will be towards their own mean rather than the general American mean. In other words, McNemar believes that *if a group of persons with extreme scores is selected on the basis of extreme first score, then regression will be towards the universe mean, but if a group of persons with extreme scores is selected on factors other than initial score, then the regression of extremes within the extreme group will be towards its own mean rather than towards the universe mean.* He admits that regression here is a descriptive rather than an explanatory concept, but although he comes near to stating the explanation, he leaves the argument at this point. McNemar's view sounds almost magical but the present paper attempts to show its essential correctness and to seek explanations. These will involve two factors, namely errors of measurement (and all that this signifies) and irregularity of growth processes.

IV. SUGGESTED EXPLANATIONS

(a) *Errors of measurement*

This concept in psychology is usually taken to cover relatively temporary factors which interfere in some way with the operation of a hypothetical 'true' intelligence (or other quality) in the test situation. The empirical basis for this assumption is the lack of a perfect correlation, already referred to, which theoretically should exist between test and immediate retest. The following remarks, therefore, are entirely confined to a consideration of the causes underlying the failure of *immediate* test-retest correlations to reach unity.

The phrase 'chance' is often used in connexion with errors of measurement;

* Thorndike (1933, 1940) believes that 0.9 is the most likely coefficient for immediate retest correlation but there is of course some variation, depending on sample, test and tester, around this value.

sometimes 'chance errors' are regarded as synonymous with the latter. It is our contention that 'chance' like 'regression to the mean' is something of a catch-phrase,* used to explain anomalies in the test situation and to dismiss unexpected facts which may 'spoil' results. Furthermore, it is felt that this much-used concept requires frequent and detailed examination, and in this connexion the late Sir Alexander Fleming made an important statement when he said: 'Never neglect any appearance or happening which seems to be out of the ordinary; more often than not it is a false alarm but it *may* be an important truth....'

It is our submission that 'errors of measurement' may involve two groups of factors, which may separately or together influence a test score, and further, that it will depend upon the nature of the test which of the two groups is the more influential.

Factor I includes such effects as alterations in mood, fatigue, boredom, illness, emotional disturbance (most of which may tend to produce errors in logical thinking), changes in motivation, misunderstanding of the instructions, practice, learning or test sophistication, failure on the part of the tester to establish rapport, to score correctly or to give the test correctly.

Factor II involves 'chance' in the narrow sense of the concept, the deliberate 'taking a chance' or guess-work by the subject.

The first group of variables may affect any test situation, the second, however, will clearly operate in some circumstances more than in others. On the whole, intelligence tests are constructed of two types of problem, the 'open' and the 'closed' (or forced choice). The 'open' type includes, for example, Wechsler's comprehension, arithmetic and vocabulary tests, where the subject is required to give an entirely individual and independent answer to an open question. The 'closed' type includes Raven's Progressive Matrices, the Mill Hill Vocabulary (Set A, Synonyms) and a number of paper and pencil group intelligence tests, where the subject is required to *select* from a limited number of responses the one which is the most appropriate. One would expect a subject to get on the average one in six correct on the Mill Hill Vocabulary (Set A) simply by spinning a coin, whereas one would not expect him to get any right on the Wechsler Vocabulary by pure guess-work. Putting it another way, a Chinese peasant, for example, with no knowledge of English whatever, would be likely to score a certain number of correct responses on the former Vocabulary test, but none on the latter. Norms, of course, take this into account.

We come now to consider how these two factors operate in the test situation and to account for the observed fact of regression of extreme scores arising from the so-called 'errors of measurement'.

Factor I. These variables tend to have effects which are normally distributed around a person's 'true' score. Clearly, however, they cannot be so distributed at the extremes of a distribution, particularly if the test has an absolute 'floor' and an absolute 'ceiling' (e.g. the Progressive Matrices). Fig. 1 suggests the effects of Factor I at the extremes.

* 'Non-homoscedasticity' is yet another catch-phrase frequently invoked to account for I.Q. variability. 'Because of the fan-shaped test-retest scattergram', it is said, 'high I.Q.'s tend to alter more than low I.Q.'s'. Obviously the fan-shaped distribution here is an effect not a cause of change, which must be sought in other factors.

Obviously, if a person has a near-zero or near-top score on initial test, then on retest Factor I errors cannot take him below zero or above the top score respectively, but can only appreciably alter his position in the inwards direction; and the greater the distance from the mean, the greater will be this effect on the average. For example, a score of 5 out of 60 has only 4 below it but the whole distribution above it so that Factor I variables have on the average a greater chance of increasing than decreasing the score.

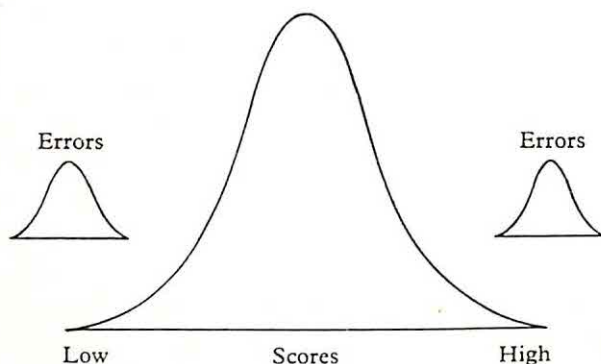


Fig. 1

Factor II. 'Taking a chance' or guessing are far less important sources of error (i.e. of departure from 'true' score) than the many aspects subsumed under Factor I, and little need be said about them. Nevertheless, like Factor I, they will tend to have effects which are normally distributed so that Fig. 1 again applies. In addition (a) they operate almost exclusively in closed tests such as Wechsler's Picture Arrangement and others already mentioned; (b) the more difficult the test item the more likely is guess-work to be used; and (c) as a corollary, the less intelligent the subject, the more likely is guess-work to be used.

The effects of Factors I and II can be summarized as follows: in tests where these play any part in successes and failures (and for Factor I this is true of most) regression of extreme scores is likely on immediate retest, solely because of errors of measurement which, providing that the test is properly standardized, are due in the main to short-term personal fluctuations. It should not, however, be inferred that regression is a simple phenomenon, since it would appear to result from the interaction of many variables. For example, practice effects are found to be greater in persons of high than low intelligence. In the former regression will occur, while in the latter, regression will result, the amount of regression from this source being greater than regression. On the other hand, since the items on most mental tests become progressively more difficult, it tends to be harder for an individual to obtain higher than lower scores; therefore, on these grounds alone, it is rather more likely that he will obtain his original score or less. It should, however, be noted that this will result in regression only for top scores since for low scores regression will occur. Furthermore, in tests which have an absolute 'floor' and 'ceiling', regression will tend to occur in persons of both high and low intelligence. The matter is further complicated by the fact that when any extreme group is selected on the basis of first test extremes, those with high I.Q.'s have sometimes gained their scores due in part to favourable Factor I

and II effects, while those of low I.Q. have sometimes gained their scores partly due to unfavourable effects which on retest will result in regression.*

Finally let us note that these effects are always behavioural primarily and statistical only secondarily; the phrase 'errors of measurement' tends to divert us from this fact because it wrongly concentrates upon the measurement rather than on the quality measured. There has been little interest in such errors because psychologists have been much more concerned with intra-individual constancy than with inconstant factors. Sir Cyril Burt (1955 and personal communication) also stresses implicitly the behavioural basis of errors of measurement. 'The word "error" tends to suggest an antithesis between a "true" measurement and a measurement that includes a superadded "error-factor". It seems... that the real antithesis is between what is *relevant* to your purpose at the time and what is irrelevant. Thus my own definition of an error of measurement would be "that part of the measurement that is attributable to factors irrelevant to the quality or quantity I want to estimate".'

(b) *Growth as a non-linear process*

We come now finally to a consideration of the progressively diminishing test-retest correlation with time. This diminution has been repeatedly demonstrated in the age range 5-20, and of course in infancy also. So far as the normal adult population is concerned, too little is known for any precise conclusions to be reached, but it is most unlikely that the correlation will diminish as rapidly once the peak of intellectual growth has been reached.

Factors affecting long-term test-retest correlations we believe to be twofold:

(1) Recovery from severe emotional disturbances which disrupt the activity of potential intelligence. This cause of I.Q. variability has been discussed by many writers and attention was drawn to it by Vernon (1955). In addition, gross deterioration precipitated by organic or functional disease such as senile dementia or schizophrenia is also relevant.

* On second testing, favourable Factors I and II effects influence low I.Q. group averages more than unfavourable, and, at the top end of the distribution, unfavourable Factor effects have greater influence than favourable. This can be demonstrated in the following manner. If there were equal numbers of persons with different I.Q.'s forming a rectangular distribution, and we were to select on first test all those with I.Q.'s between, say, 120 and 130, some would have obtained their score by favourable and others by unfavourable Factors I and II effects. Yet others would have obtained it because it was their 'true' score (i.e. score minus Factor effects). Now, for the sake of argument, suppose the maximum range of Factors I and II effects were 10 points, in any one direction then it would be possible for a person of real I.Q. 110 or 140 to fall in the 120-130 range. Since Factors I and II effects tend to be normally distributed around any point in a distribution (except at the 'floor' or 'ceiling', as already mentioned) we can state that, in our example, $X\%$ of those obtaining a score of 110 and $X\%$ of those obtaining a score of 140 will fall in the 120-130 range. If these persons were retested it would be unlikely that the same Factor effects would again favour or hinder them to the same extent and they would thus tend to approximate to their 'true' scores. The number of persons in our rectangular distribution with real I.Q.'s of 110 and 140 would be identical, therefore both groups would be equally represented in our selected sample. Thus on retest, though individuals would probably change their scores, the average score would remain approximately the same (assuming no practice effects).

Now let us consider the case of normal, as opposed to rectangular, distribution of intelligence. The same criteria will apply; a similar percentage, with 'true' scores of 110 and 140 being affected by favourable and unfavourable Factor effects respectively. Due, however, to the distribution of intelligence, more persons score 110 than 140. Therefore $X\%$ of the former represents a greater number of people than $X\%$ of the latter. Thus, if we retest persons in our respective range of which more obtained that score by favourable than unfavourable Factor effects, we must expect a lowering in average I.Q. (regression). A similar argument is relevant to the lower end of distribution.

(2) Non-linearity of growth processes which includes variation in the age when mental growth ceases.

The first factor probably operates to a rather limited extent in the general population although it is of great importance to clinical psychologists. Attention will therefore be concentrated on growth as a non-linear process.

V. PHYSICAL MEASUREMENT

In general, individual growth of all kinds tends to be non-linear with time. As Dearborn & Rothney (1941) demonstrate: 'Growth, both mental and physical, seems to be characterized not infrequently by cycles...'. The assumptions underlying the measurement of mental growth are somewhat controversial (e.g. the problem of possible inequality of units of measurement at different points in a scale, as well as the problem of an absolute zero). So we will begin by considering physical growth and revert to McNemar's argument. As we have seen, McNemar believes that the essential factor for distinguishing regression towards the universe mean from regression of the mean of the specific population lies in whether or not initial selection of subjects was on the basis of first extreme score. We argue further that errors of measurement are unlikely to be large, are probably fairly uniform throughout the ranges of physical measurement, and are therefore unlikely to affect differentially the extremes of the distributions. Thus regression of physical measurements either to the universe mean, or indeed to the mean of a specific population, is *unlikely to be due to errors of measurement but to more intrinsic factors*. (Nevertheless, physical measurements are by no means so precise and error-free as many would believe (see Thompson in Carmichael, 1954) although in general they are more accurate than mental measurements.)

The explanation appears to be quite simple and to depend on the non-linearity of growth with time. It is usually assumed that genetic factors allow a broad range of possibilities of development and the specific course of growth or point reached in an individual will, within his genetic framework, depend on environmental factors (e.g. nutrition). For example, it is unlikely that the genetics of height have altered during the last half century but there has in fact been a significant increase in mean national height. Probably under good conditions, most reach somewhere to their upper limits of potential and under bad, their lower limits. Putting it a different way, there seems to exist within persons a tendency which (a) makes extreme deviation from their own 'personal mean' a matter of low probability, and (b) the maintenance of extreme deviation for any lengthy period during growth even more improbable. The only exception to these rules is the period at which, for anyone, growth ceases; and here our 'musical chairs' or 'roundabout' theory may be helpful. A person may for some years have been fluctuating around a given extreme (with respect to others of his age-group), or indeed gradually reaching it. When growth ceases, he may be on the apogee of his fluctuation and rather like the 'odd man out' in a game of musical chairs, or a seat on a roundabout when it happens to have stopped at the peak of the sine curve it describes, be left as a rather more extreme deviant than in the past he normally has been.

We can now turn to examples which will make this clear. Fig 2 shows (a) the

average weight curve for young babies, together with the curve for one particular baby, and (b) an average and an individual height curve for one child, which illustrate in its simplest form the general principle running through this discussion.

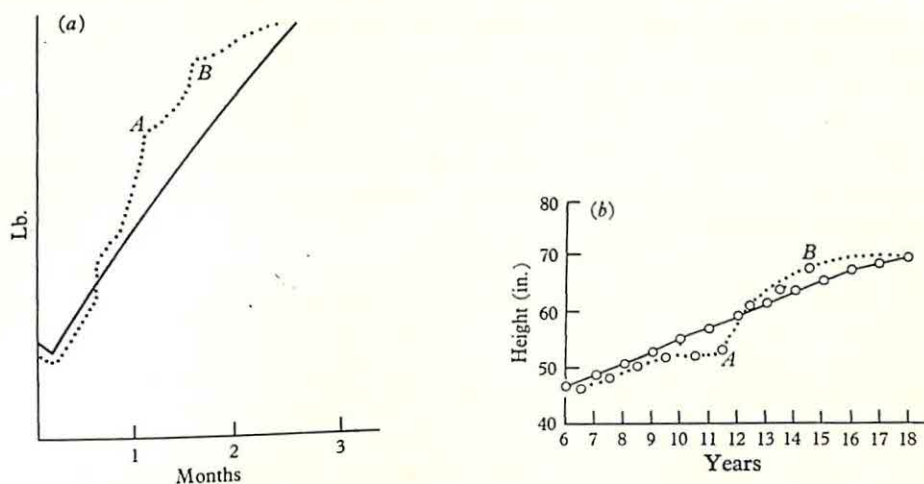


Fig. 2. (a) Average weight increase for young babies (continuous line) and the weight increase of one baby (dotted line). (b) After Thorpe (1946). The growth in height of one subject (dotted line) compared with the average (continuous line).

The baby and the child happen to show what for them are two extreme deviations at points A and B. But for every point of extreme deviation, there are many more of less extreme deviation. This illustrates McNemar's essential correctness in stating that a group selected on the basis of extreme test score will regress towards the population mean subsequently. A group of babies selected as being heavy, or children tall, for their age, would include a number who at the particular time of selection were at their points A or B. And as a group, they, like the examples in Fig. 2, would be more likely to be less extreme later or indeed earlier. But this regression is in no sense due to errors of measurement but simply because of irregularity of the growth process.

Similarly, McNemar's tall 8-year-old children were, at the time of selection, extreme deviants (for their age-group) from the population mean. Later, because of non-linearity of growth, they would tend to be relatively less extreme. The 8-year-old Swedish children, on the other hand (selected for being Swedish), would show regression of their own extreme deviants in height, but at 16 would maintain their average deviant position with respect to the general population. Both are due to inconstancy of growth increments within their respective genetic potentialities.

The same general principles are clearly at work in the two individual curves shown in Fig. 3. Once again, extreme deviation from the population mean does not occur very often and is never maintained for long.

VI. MENTAL MEASUREMENT

We are now in a position to turn to mental tests. In spite of the special problems of mental measurement, mental growth curves do not seem essentially very different from physical growth curves, and once again, exhibit variations in rate of develop-

ment in individual cases. Fig. 4 shows a variety of mental growth curves taken from various sources and samples. It will be seen that precisely the same points discussed in detail with reference to height and weight seem to apply to cognitive growth. It should be noted, however, that the analogy between physical and mental growth, though a very useful and instructive one, fails to hold at two points. The first, already mentioned, is that with physical measurements the errors are smaller and probably uniform throughout the particular range, but the second is also important, that 'with mental measurement the mere fact of having been tested affects the (later) result...' (Sir Cyril Burt, personal communication). No doubt factors such as these account for the rather greater variability of mental test results than of physical measurement.

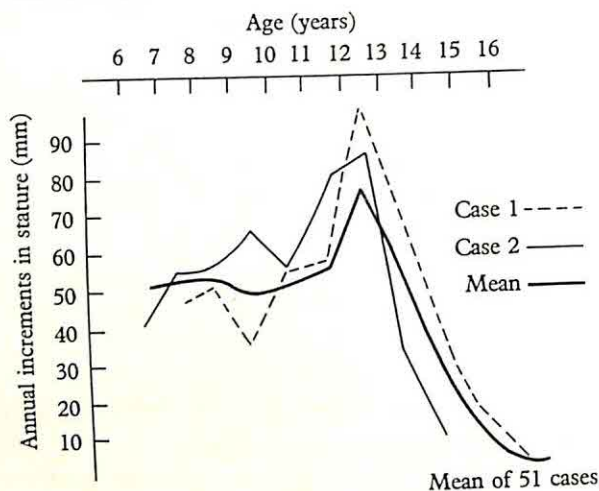


Fig. 3. After Dearborn & Rothney (1941). This shows annual increments in height of two girls compared with the mean of fifty-one subjects, chosen at random from a larger study.

We have seen that Factors I and II errors will lead to regression of extreme scores and that such changes represent fleeting, inconstant personal fluctuations. As noted earlier, regressions of this sort can best be assessed by immediate retest, and on any well standardized test (e.g. WAIS, etc.) they are unlikely to be large. But do we possess any way of establishing over and above this whether trends towards the mean are due to growth changes as opposed to Factors I and II effects? The following criteria do in fact aid this distinction, although of course there is no reason why genuine test score increment or decrement may not be superimposed on change arising from errors of measurement.

(1) Selection of a group on the basis of initial extreme mental test score will result in 'chance' regression towards the population mean. The maximal amount of such regression can be determined by immediate retest.

(2) If an extreme group be studied initially, and there is subsequently a wastage of subjects for reasons of less extreme initial test scores, then the group will show a greater than usual regression of its mean towards the population average due to errors of measurement.

(3) If a test has an immediate retest correlation of 0.9 or above, on a given population, then regression due to errors of measurement must be small. Such errors tend

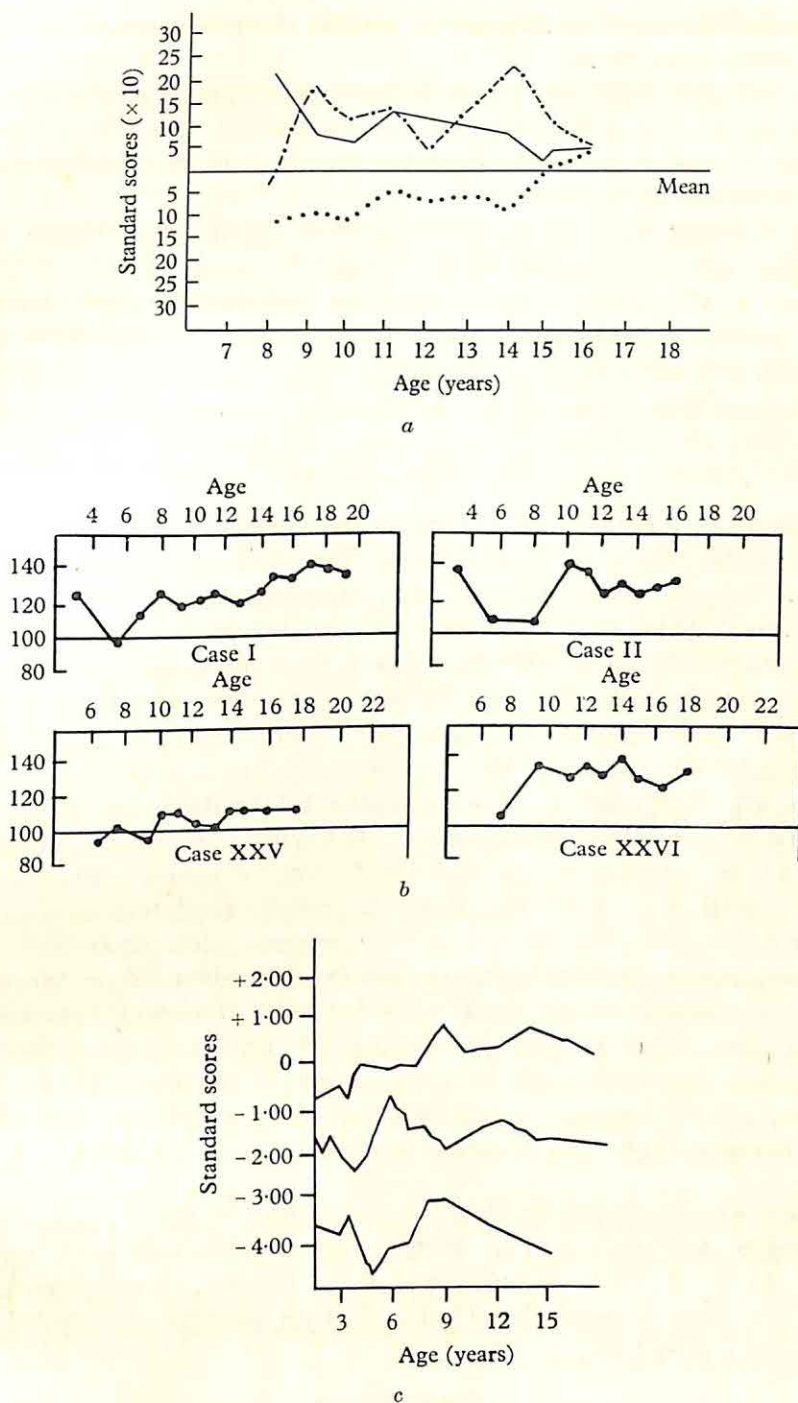


Fig. 4. (a) After Dearborn & Rothney (1941). This shows the earlier cognitive history of three subjects who happened to have reached the same point, 0.5 sigmas, above the mean of 256 girls at age 16. (b) After Hilden (1949). This shows I.Q. changes in four out of thirty cases studied. These fluctuations are probably exaggerated owing to variability of standard deviation on the Stanford-Binet; Hilden, however, also discusses the results in terms of standard scores. (c) After Honzik, Macfarlane & Allen (1948). This shows longitudinal data on three cases, and illustrates the usual finding of greater variability of I.Q. in the earlier years.

to be non-cumulative, and if as time passes greater changes occur, they must be due to something other than these.

(4) If changes occur which are linked with obviously meaningful external factors, are long-term in trend and are more or less maintained, this will militate against explanations in terms of errors of measurement which have neither constancy nor permanent cognitive significance.

(5) If initially equated groups regress to significantly different extents, then something other than errors of measurement is involved.

In summary, in this section we have tried to show that because of variations in physical and mental growth rates, there exists a tendency to regression of extreme deviants. Such regression is primarily organismic and only secondarily statistical, and is certainly neither physically nor psychologically meaningless.

VII. CONCLUSIONS

The explanations advanced in this paper to account for regression have been essentially twofold: 'errors of measurement' (Factors I and II) the extent of which can on the average be determined by immediate retest, and 'non-linearity of the growth processes'. This latter, of course, needs further explanation, which is beyond the scope of this article; most authorities would, however, agree that it is a compound of intrinsic and extrinsic factors. It is the task of experimental method to investigate the factors relevant to regression in a particular population, and methodologically this is a straightforward procedure. In conclusion, to say that 'test scores have changed because of regression' gives no greater information than to say that they have changed because they have changed. The phrase 'regression to the mean' is thus meaningless unless it is qualified by 'errors of measurement' or by 'non-linearity of growth processes', and these themselves need further qualification in a particular case if they, too, are not to become mere catch-phrases.

It may have been noted that the principles we have advanced are precisely similar to those thought to describe the data 'of biological regression' between generations. But in both cases, when considering the total population, *egression* from the mean usually balances regression, and in both, a law of central tendency, i.e. of the improbability of reaching and maintaining an extreme position with respect to the population average, seems (for a variety of reasons) to be involved.

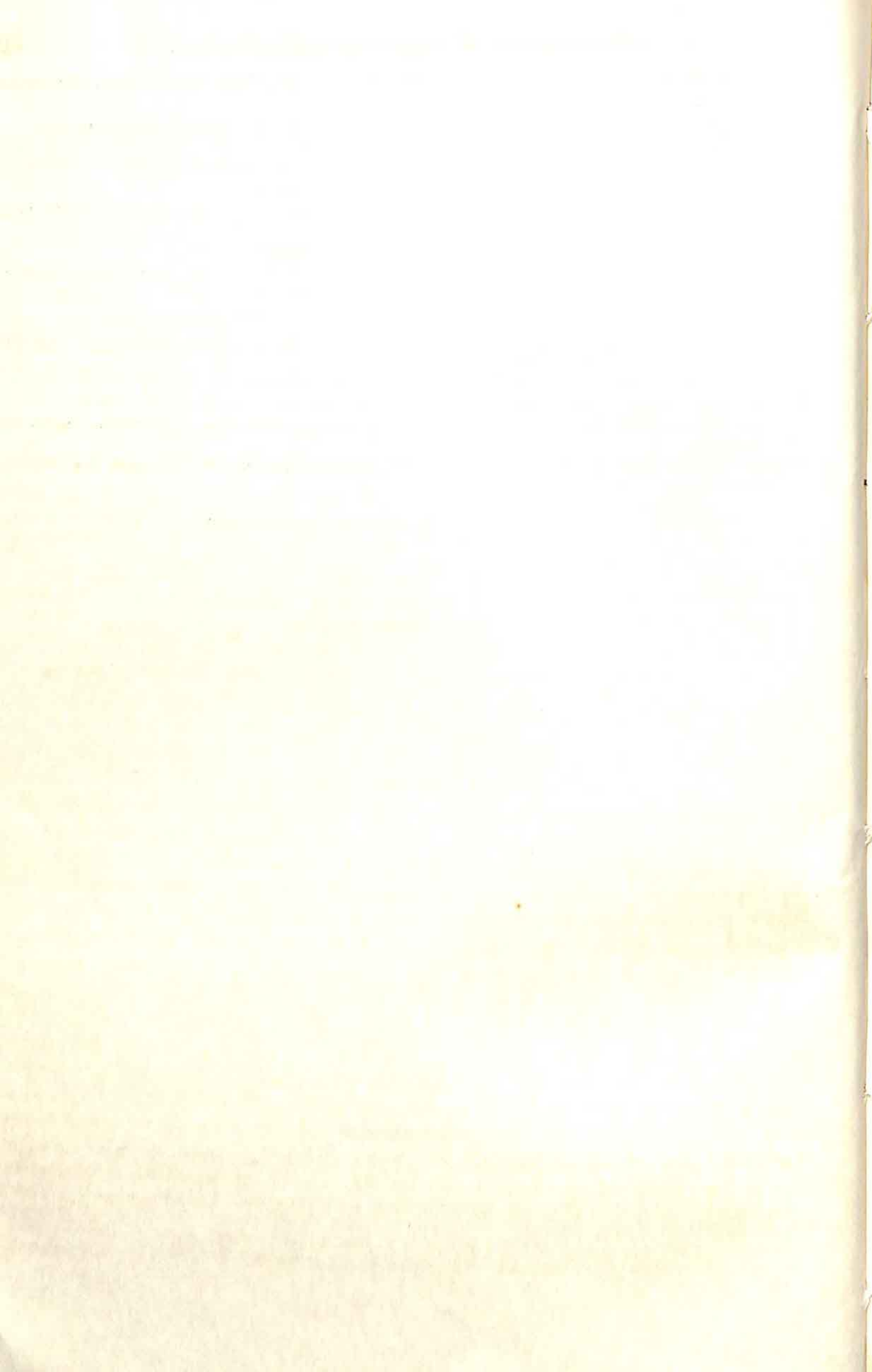
The writers wish to thank Dr Max Hamilton, who originally suggested that this paper should be written and who later made valuable criticisms, and Dr J. F. MacMahon, Dr G. W. Granger and Mr J. P. N. Phillips, all of whom made useful comments. We are particularly indebted to Sir Cyril Burt for his detailed suggestions and constructive criticisms.

REFERENCES

- BURT, C. (1955). Test reliability estimated by analysis of variance. *Brit. J. statist. Psychol.* 8, 103-18.
- CARMICHAEL, L. (Ed.) (1954). *Manual of Child Psychology*. London: Chapman and Hall.
- CLARKE, A. D. B., CLARKE, A. M. & REIMAN, S. (1958). Cognitive and social changes in the feeble-minded—three further studies. *Brit. J. Psychol.* 49, 144-57.
- CRONBACH, L. J. (1949). *Essentials of Psychological Testing*. New York: Harper.

- DEARBORN, W. F. & ROTHNEY, J. W. M. (1941) *Predicting the Child's Development*. Cambridge, Mass.: Sci.-Art.
- EYSENCK, H. J. (1953). *Uses and Abuses of Psychology*. Harmondsworth: Penguin.
- HILDEN, A. H. (1949). A longitudinal study of intellectual development. *J. Psychol.* **28**, 187–214.
- HONZIK, M. P., MACFARLANE, J. W. & ALLEN, C. (1948). The stability of mental test performance between two and eighteen years. *J. exp. Educ.* **17**, 309–24.
- MCNEMAR, Q. (1940). A critical examination of the University of Iowa studies of environmental influence upon the I.Q. *Psychol. Bull.* **37**, 63–92.
- MILES, C. C. (1954). Gifted Children. [See chap. 16 in Carmichael (1954).]
- SCHMIDT, B. G. (1946). Changes in personal, social and intellectual behaviour of children originally classified as feeble-minded. *Psychol. Monogr.* **60**.
- THOMPSON, H. (1954). Physical Growth. [See Chapter 5 in Carmichael (1954).]
- THORNDIKE, R. L. (1933). The effect of the interval between test and retest on the constancy of the I.Q. *J. educ. Psychol.* **24**, 543–9.
- THORNDIKE, R. L. (1940). 'Constancy' of the I.Q. *Psychol. Bull.* **37**, 167–86.
- THORPE, L. P. (1946). *Child Psychology and Development*. New York: Ronald Press.
- VERNON, P. E. (1951). Recent investigations of intelligence and its measurement. *Eugen. Rev.* **43**, 125–37.
- VERNON, P. E. (1955). The psychology of intelligence and *g*. *Bull. Brit. Psychol. Soc.* **26**, 1–14.

(Manuscript received 19 February 1959)



RECOGNITION OF AMBIGUOUS AND UNAMBIGUOUS VISUAL CONFIGURATIONS WITH SHORT AND LONGER EXPOSURES*

By CRAIG M. MOONEY

Defence Research Medical Laboratories, Toronto†

Three kinds of novel visual configurations—forms with incidental details, formless masses of details, forms devoid of details—comprised two parallel series for presentation in a running sequence of initial and subsequent appearances, at varying intervals. Eight of each kind of configuration comprised a series; the twenty-four, appearing initially, and subsequently, constituted a sequence of forty-eight presentations. In the sequence, two of each kind of configuration reappeared after one intervening perceptual occasion, two after five such interventions, two after ten, and two after fifteen. Subjects viewed one sequence where the configurations were tachistoscopically exposed for 0.07 sec., and the alternate sequence with configurations exposed for 5.0 sec.

A small, but significant, advantage accrued to the method of visual inspection, on both perceptual counts, under all varied experimental conditions. Apart from this, the two methods could not be differentiated on the basis of configural diversity or of lengthening of the intervals between the initial and subsequent appearances of the configurations. This want of evidence for the premise that the superiority of the method of visual inspection might be uniquely attributed to the efficacy of viewing time and scanning eye movements by virtue of contingent cognitive associations suggests that perceptual learning occurs identically with the two methods of observation and that the overall difference can be attributed to incidental procedural factors.

It is concluded on the basis of all evidence in this and earlier related studies that this simple kind of perceptual learning is not prosecuted on the basis of viewing time and scanning eye movements and that its explanation must be governed by the fact that it occurs through the single brief glance.

I. INTRODUCTION

Two previous studies (Mooney, 1958, 1959) addressed the question: are eye movements important in learning to recognize novel visual configurations? They were a logical consequence of earlier studies (Mooney, 1956, 1957*a, b*) of *closure* (perception based on incomplete representations of familiar objects) where it was found that visual inspection was no more effective than the single brief glance. They were designed to question the importance attached to eye movements in perception and perceptual learning by Hebb (1949), Gibson (1950), Lashley (1954), and others.

It was assumed in the two studies that the recognition of meaningless configurations on a second occasion of experience with them is evidence of a simple kind of perceptual learning. It was then proposed that if eye movements are essential to, or do facilitate, the 'grounding' of memories (as engrams, cell assemblies—however postulated) or their subsequent arousal, the recognition of meaningless configurations should be markedly more effective when they are initially observed by visual inspection rather than by a single brief glance. Numerous configurations were shown a first time and subsequently presented intermixed with other configurations appearing for a first time, under two methods of observation—a period of visual inspection, and a

* *DRML Report* no. 142-3, project no. D 77-94-35-29 (HR no. 180).

† Now with the Canadian Joint Staff, London, England.

single brief glance. *Ss* were required to distinguish those seen before from those appearing for the first time. A series of factorially designed experiments permitted comparison of the two methods of observation with different kinds of configurations.

In the first study the configurations were supposed to be meaningless, and were of two kinds—structured and amorphous. Performance proved to be at a high level, unaffected by method of observation or kind of configuration. The premise of the second study was that shape-like or quasi-meaningful configurations might provide an associative basis that could be more effectively exploited by visual inspection than by the single brief glance; accordingly, the configurations were inkblots of two kinds, symmetrical and non-symmetrical. Again, performance under the two methods of observation was at a high level with, however, a slight but significant advantage accruing to the method of visual inspection with both symmetrical and non-symmetrical configurations. This outcome was not inconsistent with the premise; nevertheless, in view of all the findings, the null hypothesis (that visual inspection would be no more effective than the single brief glance) could not be confidently rejected.

The present study was undertaken not only to check the above findings but to provide a more rigorous test of the premise in the second study—namely: that visual inspection should be most effective in learning to recognize configurations most amenable to cognitive constructions or associations.

It seemed that the configurations hitherto used (leaving aside symmetrical ones) were of three types: (a) ambiguously structured forms with incidental details; (b) formless assemblies of incidental details; and (c) unambiguously structured forms devoid of incidental details. It was surmized that the first of these offered the greatest scope for cognitive associations and the third the least. On the presumptions that prolonged viewing time and eye movements would be more conducive to cognitive associations than the single brief glance and that the memories for configurations so construed should be more persistent than those acquired by the single brief glance, two experimental propositions were made. The first was that visual inspection would be markedly more effective than the single brief glance in learning to recognize configurations of the first kind, possibly slightly more effective with configurations of the second kind, and no more effective with configurations of the third kind. The second was that any decline in performance due to a lengthening of the intervals between the initial and subsequent occasions of experience with the configurations should be markedly less with the method of visual inspection than the method of the single brief glance.

II. METHOD

The three types of items were mounted on 35 mm. slides, in circular frames. They were categorized, respectively, as FD (forms with details), DD (details only) and FF (forms only.) Examples are shown in Fig. 1. These were casually sorted into two groups of test-items, each group being composed of eight of each of the three kinds of configurations. The twenty-four items in a group were arranged for a sequence of initial and subsequent presentations (first and second appearances)—making forty-eight presentations in all. This arrangement was such that: of each of the three kinds of configurations two would be presented with one other configuration appearing (as an initial or subsequent presentation in its own right) between their initial and subsequent appearances; two with five other presentations so intervening; two with ten such interventions; and two with fifteen.

A projector with a variable-speed light-interruptor permitted exposure of the configurations in a dark room for brief or lengthy periods as a display 20 in. in diameter on a screen 6 ft. in front of *S*. The *S* was tested with one group of items at one exposure and with the other group at a second exposure—the exposure being 5.0 sec. or 0.07 sec. (This latter exposure being brief enough to prevent a second fixation.) To anticipate the 5.0 sec. exposure *S* fixated the centre of the circle of light in which, at about 10 sec. intervals, a configuration appeared for 5.0 sec. To anticipate the short exposure, *S* fixated the centre of the circle of light—appearing in successive 0.07 sec. flashes at 1.0 sec. intervals—where, after approximately ten such flashes, the configuration appeared for 0.07 sec. (The two procedures were thus technically equated only for time elapsing—approximately 10 sec.—between cessation of one presentation and onset of the next.) Each test took about 10 min., with a recess of 5 min. between.

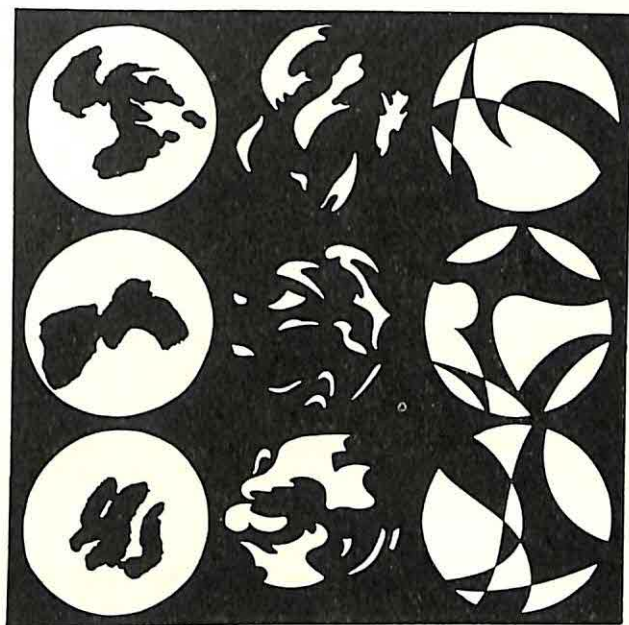


Fig. 1. Examples of meaningless configurations: left to right, FD (form & detail), DD (details only), FF (form only.)

The *S* was told that he would be viewing meaningless shapes each of which would appear a first time and reappear sooner or later in an indefinitely long series of such presentations; that he was to 'look hard' at each as it appeared in order to be able to recognize it later; that he was to indicate whether he recognized a configuration by calling 'Yes' or 'No' after each presentation, starting with the first one in the series; that by 'Yes' he would mean it was one he had seen before and that by 'No' he would mean he had not seen it before.

Subjects were twenty-four young males (aged 18–22 years, the mean being 19), applying for enlistment at an Army recruiting depot. They were informed that the tests measured visual speed and memory. They were permitted to assume that the tests were an integral part of the routine military examination.

A factorial design was employed accounting for the variables: item groups (two), order of group presentation (first, second), exposures (two), kinds of configurations (three), and appearance intervals (four). Since the types and appearance intervals of the configurations were invariant within the sequence of presentations within the item groups, replication was based on the first three *a/m* variables $2 \times 2 \times 2$. Thus a threefold replication called for the present 24 *Ss*.

Scoring was in terms of errors. An erroneous 'No' response (failure to recognize a configuration on its second appearance) constituted a non-recognition; an erroneous 'Yes' response (wrongly 'recognizing' a configuration on its first appearance) constituted a false recognition.

III. RESULTS

The data of main interest are set out in Table 1, and the appropriate analyses of variance are shown in Table 2. The significant findings are displayed in Figs. 2, 3 and 4.

Table 1. *Non-recognitions of diverse novel configurations at various intervals and exposures*

Exposure	Kind of configuration	Presentations intervening between initial and subsequent appearance				Total
		1	5	10	15	
0.07 sec.	FD	11*	9	9	19	48 (49)†
	DD	12	13	19	16	60 (59)
	FF	12	10	14	21	57 (79)
	Total	35	32	42	56	165 (197)
5.0 sec.	FD	9	1	5	21	36 (25)
	DD	8	4	11	17	40 (42)
	FF	5	4	11	15	35 (65)
	Total	22	9	27	53	111 (132)
Total		57	41	69	109	276 (319)

* Total possible per cell, 48—based on 24 *Ss* and 2 configurations per interval.

† In brackets, false recognitions of configurations on first appearance—based on 24 *Ss* and 8 configurations.

Table 2. *Analyses of variance for non-recognitions and false recognitions of diverse configurations at different intervals and exposures*

Source	D.F.		M.S.		F	
	Non-R	False	Non-R	False	Non-R	False
Item groups (<i>S</i>)	1	1	0.007	3.673	—	—
Group order (<i>O</i>)	1	1	0.028	2.506	—	—
Exposures (<i>E</i>)	1	1	5.062	21.006	13.64*	10.69*
Item types (<i>C</i>)	2	2	0.333	25.965	—	13.21*
Appearance intervals (<i>I</i>)	3	—	5.852	—	15.77*	—
Interactions:						
1st order	24	9	0.426	2.396	—	—
2nd order	34	7	0.361	1.018	—	—
3rd order	23	2	0.177	0.340	—	—
4th order	6	—	0.305	—	—	—
Within cells	480	120	0.371	1.965	—	—
Total	575	143	0.399	2.408	—	—

* $P = 0.01$.

Looking at Tables 1 and 2, there were no significant first order or lower interactions—hence, they have been consolidated. It will be noted that the method of visual inspection was slightly, but consistently and significantly, superior to the single brief glance in minimizing both non-recognitions and false recognitions.

Non-recognitions (Fig. 2) were unrelated to the type of configuration employed, running at about 20 % by the method of visual inspection and 30 % by the method of the single brief glance. False recognitions (Fig. 3) were markedly affected by the type of configuration with both methods of observation—the percentages being,

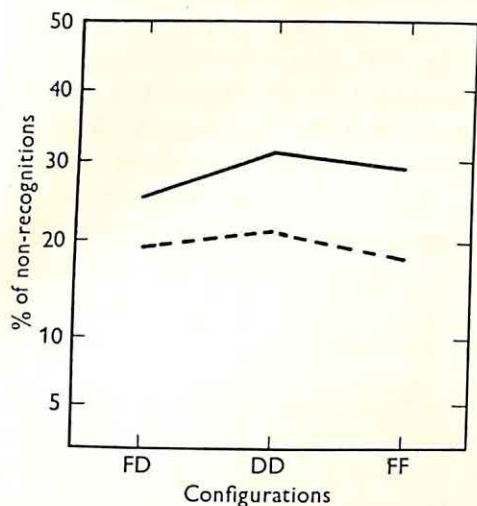


Fig. 2

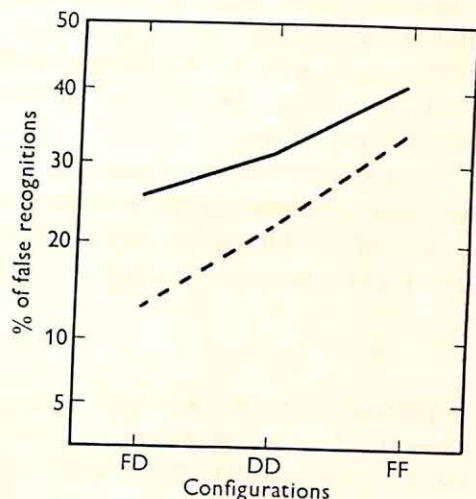


Fig. 3

Fig. 2. Non-recognitions of configurations on second appearance in terms of kinds of configurations and exposures employed. Exposure: —, 0.7 sec.; - - - 5.00 sec.

Fig. 3. False recognitions of configurations on first appearance in terms of kinds of configurations and exposures employed. Exposure: —, 0.07 sec.; - - - , 5.00 sec.

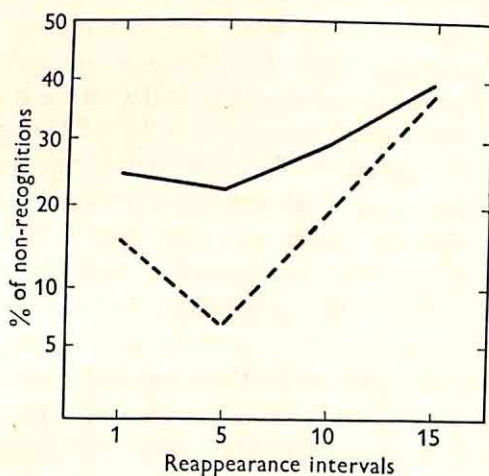


Fig. 4. Non-recognitions of configurations on second appearance in terms of intervals between first and second appearances. Exposure: —, 0.07 sec.; - - - 5.00 sec.

for FD, DD and FF configurations respectively 13, 22, and 34 with visual inspection, and 25, 31, and 41 with the single brief glance.

Non-recognitions increased as a function of the number of perceptual occasions intervening between the initial and the subsequent appearances of the configurations, as shown in Fig. 4. With an interval of fifteen occasions, visual inspection was no more

effective than the single brief glance—non-recognitions approaching 40% in both cases (presumably, if extrapolated, approaching the chance level with an interval of eighteen or twenty intervening occasions.) This change occurred independently of the kinds of configurations employed.

Relating these findings to the initial propositions. The assumption of inherent differences in the three types of configurations is borne out by the systematic differences in false recognitions (see Fig. 3). But the first experimental proposition (that visual inspection should be most markedly superior to the single brief glance with FD configurations and least so with FF) is not borne out (see Fig. 2). Neither is the second proposition (see Fig. 4)—namely, that with increasing lengths of time and numbers of interposed presentations between first and second appearances of particular configurations any decline in performance should be markedly less with visual inspection than with the single brief glance.

IV. DISCUSSION

The initial question has been: is perceptual learning, of the kind incidental to the recognition of novel visual configurations on a second occasion of experience with them, more effectively accomplished by visual inspection than by a single brief glance? The present findings, like those of the previous study (1959), reveal a high level of accomplishment by both methods in distinguishing novel configurations on their first appearance and in recognizing them on their second appearance, with a slight, but significant, advantage accruing to the methods of visual inspection.

The subsequent question has been: is this slight advantage in favour of visual inspection an incidental procedural consequence, or is visual inspection an intrinsically superior mode of perceptual acquisition? The latter possibility, it was assumed, would reside in the fuller development of cognitive associations through the agency of viewing time and scanning eye movements, with the collateral assumption that this would enhance the subsequent recognizability of the configurations. It was reasoned, in brief, that the method of visual inspection might prove maximally advantageous with configurations that provided greatest associative possibilities. But the findings are not consistent with such reasoning. The advantage possessed by visual inspection proves not to be uniquely associated with configural properties—with, for example, ambiguity, as in the present study, or symmetry or structural coherence, as in the earlier studies. The perceptual accomplishments under the two methods of observation do not diverge in the expected directions. The failure of the two methods to split apart on the basis of configural diversification, or temporal extension, means that the superiority of the method of visual inspection cannot be accounted for by the compound thesis originally proposed. Either the factors of viewing time and scanning eye movements are not salient to cognitive associations, or these latter are not relevant to this kind of perceptual learning.

Since the logical brief for visual inspection as an essentially different method of perceptual learning cannot be substantiated, its slight absolute advantage may be attributed to incidental experimental factors. Since no objective recordings of eye movements were made, there is the possibility that they did not occur during the 5.0 sec. exposures. But this is unlikely. *Ss* invariably testified that they did search

or study the displays during long exposures; moreover, experience with this kind of testing situation offers no grounds for supposing that *Ss* would maintain a single fixation during the 5.0 sec. periods. Another possible factor is that of slight differences in timing between the two methods of presentation. For example, while 10 sec. intervals were maintained between cessation of one presentation and onset of the next, with both methods, the intervals were uninterrupted between long exposures and broken into ten successive flashes between brief exposures. Also, by holding the intervals constant differences in elapsed time were introduced; for example, with the second presentation of any one item coming ten items after its first presentation, the total period of elapsed time would be about 110 sec. with the brief exposures and about 160 sec. with the long ones—and, with fifteen items interspersed, 160 sec. and 230 sec., respectively. Thus, so far as such absolute differences in elapsed times—apart from the number of interspersed items—may be critically important, then, as these latter were multiplied, the long-exposure method would be at an increasing disadvantage. But this becomes tenuous. A simpler and more likely explanation may be that performance with the brief exposures was slightly handicapped throughout by the method of tachistoscopic presentation employed or by some inhibition of subjective certainty through the disconcerting effects of a relatively unnatural method of observation.

The general conclusion, in the light of all data from these several studies, would seem to be that the learning process entailed in initial familiarization with novel visual configurations is independent of viewing time and scanning eye movements, and that explanations of such learning must be constrained, therefore, by the fact that it occurs through the agency of an initial, single brief glance.

REFERENCES

- GIBSON, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton Mifflin.
- HEBB, D. O. (1949). *The Organization of Behavior*. New York: Wiley.
- LASHLEY, K. S. (1954). Dynamic processes in perception. In *Brain Mechanism and Consciousness: A Symposium*. Springfield: Thomas.
- MOONEY, C. M. (1956). Closure with negative after-images under flickering light. *Canad. J. Psychol.* **10**, 191-9.
- MOONEY, C. M. (1957a). Closure as affected by viewing time and multiple visual fixations. *Canad. J. Psychol.* **11**, 21-8.
- MOONEY, C. M. (1957b). Closure as affected by configural clarity and contextual consistency. *Canad. J. Psychol.* **11**, 80-8.
- MOONEY, C. M. (1958). Recognition of novel visual configurations with and without eye movements. *J. Exp. Psychol.* **56**, 133-8.
- MOONEY, C. M. (1959). Recognition of symmetrical and non-symmetrical ink-blots with and without eye movements. *Canad. J. Psychol.* **13**, 11-9.

(Manuscript received 16 March 1959)

THE OPTIMAL PERCEPTUAL LOAD IN A PACED AUDITORY INSPECTION TASK

By E. C. POULTON

Medical Research Council, Applied Psychology Research Unit, Cambridge

A paced auditory inspection task was presented at four levels of perceptual load and three rates of critical signals. The aim was to determine whether there was an optimal range of perceptual load, over which deterioration in performance with time would be minimal. Beyond this optimal range, where the load was either too great (soon producing 'mental fatigue'), or too small (soon producing 'boredom'), greater deterioration was to be expected. The experimental results supported the hypothesis.

The task with the smallest perceptual load presented critical signals only once about every 4 mins. on average. This vigilance task gave significantly more errors than a task with a greater perceptual load and a higher rate of critical signals. Conditions were designed to maintain motivation in both cases. Thus the poorer performance is unlikely to have been due simply to inadequate motivation.

I. INTRODUCTION

(i) *Changes in perceptual load*

When performance in an inspection task deteriorates with time (Wyatt & Langdon, 1932, Fig. 10), psychologists have two common-sense explanations available. If the operator's perceptual mechanism is heavily loaded, we can say that he must have become mentally fatigued. When the perceptual load is light, we can say that he must have become 'bored'. If these common-sense explanations are valid, there should be a middle range of perceptual load in inspection tasks, over which performance deteriorates less with time than it does when the load is too great or too small. In the case of an industrial paced inspection task lasting 8 hr. every weekday, we should predict a smaller increase in the proportion of errors towards the end of a spell of work for a task with a medium perceptual load. In the case of a brief session such as was used in the present experiment, we might expect no detectable increase in errors on the first performance of a task with a medium perceptual load, as contrasted with a detectable increase in tasks with greater and smaller loads.

It is possible for deterioration during an experimental session to be masked by learning during the session (Thorndike, 1912). In order to test for a hidden deterioration, it is necessary to conduct a second session under identical conditions. The score at the start of the second session, when the subject is fresh, gives an indication of how he would have performed at the end of the previous session if he had not become 'mentally fatigued' or 'bored'. However, the hypothesis of no detectable increase in errors for a task with a medium perceptual load, may not hold for experimental sessions other than the first. With repetition brief experimental sessions may come to resemble more closely the typical industrial situation, in which some deterioration in performance with time is likely to occur at all levels of perceptual load.

In the present experiment perceptual load and rate of critical signals sometimes varied concurrently. We have therefore to be careful not to attribute to a reduction in perceptual load an effect actually due to a reduced rate of critical signals. The

experiment was *not* concerned with subjective feelings during or after a spell of work. The aim was simply to relate perceptual load to decrement in performance. We wished to discover whether this function bears any relationship to common-sense explanations.

Previous work on perceptual load has studied only the upper limit (Conrad, 1951; Mackworth & Mackworth, 1956, 1958; Wagner, Fitts & Noble, 1954). It appears to have been assumed that the smaller the perceptual load the better. Studies of performance associated with reduced vigilance or 'boredom' have emphasized the low rate of critical signals rather than the small perceptual load. The only possible exception is Jerison & Wallis (1957), who compared monitoring one 'clock' with monitoring three.

(ii) *Motivation in vigilance tasks*

There have been a number of studies in the laboratory of inspection tasks involving a reasonably difficult discrimination, a small perceptual load, and relatively infrequent signals requiring any action. In these vigilance tasks performance has often been found to deteriorate fairly rapidly. The deterioration has most frequently been measured in terms of the proportion of failures to respond within a specified time limit, following Mackworth's original experiments (1948, 1950).

Sometimes deterioration with time in vigilance tasks is accompanied by, or replaced by, a low average level of performance. In this case performance can often be improved by some relatively minor modification of the task, e.g. by increasing the rate of critical signals (Jenkins, 1958; Deese, 1955; Bowen, 1956; Bakan, 1959; Kappauf & Powe, 1959; Garvey, Taylor & Newlin, 1959).

In the typical vigilance experiment the subject works by himself at a laboratory task which may appear rather pointless. In many cases he is never told how well he has done, nor given any particular encouragement to try hard. In these circumstances it is not surprising that his efficiency is not always well maintained. The importance of motivation is shown by Mackworth's telephone-message experiment (Mackworth, 1950, pp. 22-24). Mackworth presented a brief personal telephone message to the subject in the middle of the experiment, when performance appeared to be at low ebb. The message raised the level of performance for the next $\frac{1}{2}$ hr. to that of a fresh subject at the start of the experiment.

A method of maintaining motivation might prevent, or at least mitigate, the lowered level of performance which is otherwise to be expected. In the present experiment the experimenter remained in the room, and personally recorded the subject's responses. He encouraged the subject to do his best, and gave him at the end some uninformative appreciation of his performance. It was hoped that this might prove an effective method of maintaining an adequate level of vigilance.

II. METHOD

(i) *Experimental subjects*

The 96 subjects were all naval ratings, aged between 16 and 28 years, with a median age of 21. Their scores on intelligence test AH 4 (Heim, 1955) ranged from 47 to 108, with a median of 71.

(ii) *Experimental material*

Strings of eight numerals were recorded on magnetic tape, one string every 4.0 sec. The numerals in a string were read staccato in time to a metronome at a rate of three numerals per sec. Each string was identical with the previous string, except for one changed numeral. The first eight changes in the practice material are shown in Table 1. The last string in the table is quite different from the initial string, since each new numeral remains in subsequent strings until it is replaced by another new numeral. The material thus differs from that in a typical industrial inspection task. In industry the standard against which each successive item is judged remains constant.

Table 1. *Example of experimental material*

Serial position	1	2	3	4	5	6	7	8
Initial string	... 2	5	4	8	7	1	5	7
Practice string no.								
1	2	5	3	8	7	1	5	7
2	2	5	3	6	7	1	5	7
3	2	5	3	6	7	1	5	2
4	2	8	3	6	7	1	5	2
5	2	8	3	6	7	6	5	2
6	2	8	3	6	7	6	1	2
7	2	8	3	6	7	4	1	2
8	5	8	3	6	7	4	1	2

Note: The new numeral introduced in each string is in bolder type.

The changes in successive strings were intended to appear random, but the following restrictions were imposed upon randomness:

(a) Only the numerals 1 to 8 were used.

(b) In a set of 64 consecutive strings, each of these eight numerals appeared as a change once in each of the eight serial positions.

(c) Two consecutive new numerals could not be the same, nor appear in the same serial position. In addition, the eight first appearances of each numeral in a set of 64 strings were spaced out so that all eight could not appear by chance in a small section of the set. Similarly, the eight changes in each serial position were also spaced out within the set of 64 strings.

(d) The same numeral could not appear more than twice in a string of eight numerals. Strings could not contain easily remembered combinations such as: 22, 53 followed later in the string by another 53, 135 or 642.

(e) The string at the end of the second set of 64 consecutive strings was identical with the initial string which preceded the start of the first set of 64. Thus the total of 128 strings formed a closed loop. The complete test material contained four identical copies of this closed loop, which followed each other without a break.

(f) The appearance of a new 6 in the sixth serial position was arranged so that this critical signal was separated from an identical signal by alternately 92 and 34 other signals.

These restrictions upon randomness reduced the new information presented in each string to about 4.7 bits. It also reduced the standard deviation of the size of the intervals between the first appearances of 6's from an expected mean lying between 7.3 and 8.2 strings at the 5% limit of confidence to 3.7. The standard deviation of the intervals between the first appearances of new numerals in the sixth serial position was reduced to 6.1 strings.

The complete tape record was played to the subject at a comfortable loudness, and took rather under 34 min.

(iii) *Experimental conditions*

The main conditions are listed in Table 2.

C1 *all new* had the greatest perceptual load. The subject had to call out the new numeral immediately after each string had been presented. He had always to search the eight serial

positions (Table 2, col. 2), and to decide which of eight possible numerals was the new one (col. 3). He did not have to specify the serial position (col. 4). This condition also had the highest rate of critical signals, since every string contained a new numeral. A critical response was therefore required every 4.0 sec. (col. 5).

Table 2. *Main experimental conditions*

Conditions (in order of perceptual load)	Perceptual load				Time between critical signals (average in sec.)	No. of subjects doing condition:	
	In signal		In response			At least twice	Four times
	No. of positions to search	No. of numerals to listen for	Position must be correct	Numeral must be correct			
C1 all new	8	8	No	Yes	4	12	—
C2 only position 6	1	8	Yes	Yes	32	48	24
C3 only 6's	8	1	No	Yes	32 or 16	24	—
C4 only new 6's in position 6	1	1	Yes	Yes	256	12	—

C2 only position 6 had the next largest perceptual load. There were two versions. *C2a new position 6* involved calling out the new numeral if it was in serial position 6, and 'no' if it was in any other serial position. The subject had to search only in the sixth serial position. On the one in eight occasions upon which the new numeral appeared in this position, he had to decide which of the eight possible numerals it was. Both the numeral and the serial position had to be correct. Critical signals occurred on average every 32 sec. In the *C2b all position 6* version the subject had simply to call out after each string the numeral in the sixth serial position, whether it was a new numeral or not. The perceptual load and the rate of critical signals were the same as in *C2a new position 6*.

C3 only 6's came next in order of perceptual load. Again there were two versions. *C3a new 6's* involved calling '6' whenever the new numeral was a 6, and 'no' whenever the new numeral was any other. The subject had to search all eight serial positions, but he had only to decide whether the new numeral was a 6 or not. He did not have to specify its serial position. Thus the perceptual load was smaller than in the *C2 only position 6*, in which both the numeral and the serial position had to be correct. On average one signal in eight was critical, i.e. one every 32 sec.

In the *C3b count 6's* version the subject had simply to call out the number of 6's in each string of numerals. The perceptual load was the same as in *C3a new 6's*. The rate of critical signals was twice as great, since the subject had to change his response when a 6 disappeared as well as when a new 6 appeared.

C4 only new 6's in position 6 was a vigilance condition. It had the smallest perceptual load, and critical signals occurred on average only once every 4 min. 16 sec. The subject had to call '6' when the new numeral was a 6, provided it occupied the sixth serial position. If the new numeral was not a 6, or if it was a 6 but occupied any serial position other than the sixth, he had to say 'no'. Thus the subject had to listen either for new 6's, or for new numerals in the sixth serial position. Presumably he would make whichever of the two discriminations happened to be the easier. Once every 8 signals on average, the correct answer to this discrimination would be 'yes'. On this occasion he had then to make the alternative discrimination also. A new 6 appeared in the sixth serial position on average once every 64 signals.

(iv) *Experimental design and procedure*

The subjects were tested almost always in the afternoon. They came to a practice session generally in groups of three or four. First, they were taught the procedure for C1 all new. A few short practices on altogether 80 strings of numerals ensured that they knew what to do. They were then given a *practice test* on this condition, lasting about 5.5 min., for 80 consecutive strings. In this session the subjects had to write down their answers. After each practice they corrected them themselves, and announced their scores to the experimenter so that the other subjects could hear.

In the experimental sessions each subject was tested separately. He called out his answers to the experimenter, who was seated opposite him behind a screen. A separate group of twelve subjects performed each of the six subconditions twice in the same week. The subject was allocated to a condition according to his score on the practice test, so that each condition was performed by the same number of good, medium, and poor scorers. In addition, the order of the conditions for the first experimental session was chosen so that the method of performance of a subject in a later session would not be influenced by tips received from a subject already tested.

At the start of the first experimental session each subject was reminded of his score on the practice test, and of the scores of his colleagues. He was taught the new procedure, and was then allowed to practise it on 40 consecutive strings of numerals. This practice was repeated if necessary until the experimenter was fairly certain from the subject's performance that he knew what he had to do. The greatest number of repetitions required was four. This was on C2b all position 6.

At the start of the second experimental session the subject was given some encouraging, but not informative, indication of his general level of performance in the previous session, e.g. 'you did jolly well', or 'you didn't do too badly'. He was encouraged to do still better in the present session.

After separate groups of twelve subjects had performed each of the six subconditions twice, two additional groups of twelve subjects each were given the two versions of C2 only position 6 four times in the same fortnight (see Table 2). The scores of these two groups on the practice test were distributed similarly to the scores of the other six groups of subjects.

(v) *Scoring and calculations*

The headings of the columns in Table 3 show the main ways in which the experimental conditions were scored. All conditions were scored upon responses to new 6's in position 6. C3 only 6's and C1 all new were also scored upon responses to all new 6's. C2 only position 6 and C1 all new were scored upon responses to all new numerals in position 6. C1 all new was scored upon responses to all new numerals. C3b count 6's was also scored upon responses to the disappearance of a 6, but these responses are only considered in Result (iii). Fig. 1 uses the data from the maximum number of critical signals in each condition. The only exception is the omission of responses to the disappearance of a 6 in C3b count 6's. The ways in which different kinds of errors have been classified are described in Result (iii).

For each subject the trend in errors for the four quarters of each experimental session was fitted by an orthogonal polynomial of degree 1 (Fisher & Yates, 1948, Table 23). If the mean slope for any condition was significantly different from zero, this was taken to indicate deterioration or learning during the session. From the individual polynomials, the expected mean error was calculated for an imaginary 'fifth quarter'. If the mean error for the first quarter of the subsequent session was significantly smaller than this, it was assumed to demonstrate a decrement in performance during the previous session which was concealed by learning (Thorndike, 1912). In these and other statistical tests t was generally used. Where a t -test was inappropriate owing to the distribution of scores, a simple binomial test or a median test using χ^2 was substituted. Special statistical tests are described in the text. Two-tailed tests have always been used except where stated.

III. RESULTS

(i) *Relative difficulty of experimental conditions*

The means and standard deviations (s.d.'s) for each experimental period or day are given in Table 3. To make the different methods of scoring comparable, the raw means have been converted into errors per 100 critical signals scored. The s.d.'s represent the variability between subjects. The two versions of C2 only position 6 gave almost identical results, and have therefore been combined. The two versions of C3 only 6's have been combined for the same reason.

All the four main experimental conditions can be compared only upon the critical signals which were common to them all: new 6's in position 6 (last column in Table 3).

Each subject was only presented with eight of these signals per day. A greater number of critical signals was common to C2 only position 6 and C1 all new (middle column), and to C3 only 6's and C1 all new (one from last column). The table shows that in general the order of difficulty corresponded to the order of perceptual load, with C1 all new significantly the most difficult condition. The one exception to this rule was C4 only new 6's in position 6, which was significantly worse than C3 only 6's, although the perceptual load was lighter. C4 only new 6's in position 6 had a very low rate of critical signals (see Table 2).

Table 3. *Errors per 100 critical signals for the main experimental conditions*

Conditions (in order of perceptual load)	Critical signals scored							
	All new numerals		New numerals in position 6		New 6's		New 6's in position 6	
	Mean	S.D.	Mean*	S.D.	Mean†	S.D.	Mean‡	S.D.
	Day 1							
C1 all new	38.6	15.5	49.6	24.0	36.6	47.6	48.0	35.4
C2 only position 6	—	—	25.2	20.6	—	—	20.6	20.8
C3 only 6's	—	—	—	—	9.6	6.82	1.6	4.14
C4 only new 6's in position 6	—	—	—	—	—	—	17.7	20.8
Day 2								
C1 all new	31.5§	12.7	42.1§	19.8	32.7	16.1	34.4§	22.8
C2 only position 6	—	—	9.2§	13.1	—	—	7.3§	11.9
C3 only 6's	—	—	—	—	5.1§	4.28	1.0	3.45
C4 only new 6's in position 6	—	—	—	—	—	—	7.3	14.9
Day 3								
C2 only position 6	—	—	3.8§	6.30	—	—	1.6§	4.13
Day 4								
C2 only position 6	—	—	1.6§	2.10	—	—	1.0	3.45

* C1 all new, different from C2 only position 6 at the 0.001 level on both days 1 and 2.

† C1 all new, different from C3 only 6's at the 0.001 level on both days 1 and 2.

‡ Day 1. C3 only 6's, different from all others at the 0.02 level or better. None of the other differences is significant.

Day 2. C3 only 6's, different from C2 only position 6 at the 0.05 level, but not different from C4 only new 6's in position 6. C1 all new, different from all others at the 0.02 level or better.

§ Improvement over previous day significant at the 0.05 level with a one-tailed test or better.

(ii) *Trends between and within days*

Comparing days, Table 3 shows that practice significantly improved performance. The only exception was C4 only new 6's in position 6. This was probably due to the small amount of data available in this condition. Only nine subjects made any errors at all.

The four quarters of each experimental period are directly comparable to each other, since the experimental material was identical. Fig. 1 shows the trends between quarters for the different experimental conditions.

In C1 all new there were no significant differences between quarters on day 1 ($P > 0.05$), and there was also no significant trend. However, the average error in the first quarter of day 2 was significantly less than that predicted from the individual trends in performance within day 1 ($P = 0.01$ —see Scoring and Calculations). This suggests that there was a decrement in performance towards the end of day 1, but

that it was masked by learning (Thorndike, 1912). The hidden decrement could be attributed to 'mental fatigue', since this condition involved a considerable perceptual load.

The individual trends in errors for the four quarters of day 2 were not significant ($P = 0.13$), although the difference between the combined first two and last two quarters was just significant ($P = 0.05$). On day 1 the first 2.0 min. of the first quarter produced 20% less errors than the corresponding first 2.0 min. of any of the subsequent quarters ($P = 0.05$ or better). The corresponding figure for day 2 was 30% ($P = 0.02$ or better). There were no other significant trends between corresponding periods within quarters.

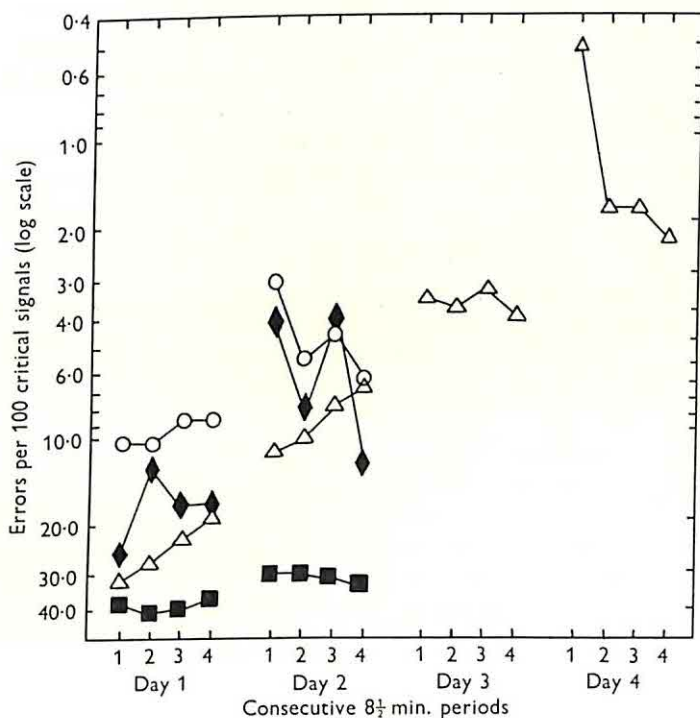


Fig. 1. Trends within and between days for the four main experimental conditions. Means are based upon the following number of observations: ■, C1 all new: 128 responses from each of twelve subjects. △, C2 only position 6: sixteen responses from each of forty-eight subjects on days 1 and 2, and from each of twenty-four subjects on days 3 and 4. ○, C3 only 6's: sixteen responses from each of twenty-four subjects. ◆, C4 only new 6's in position 6: two responses from each of twelve subjects.

C3 only 6's also showed no significant trend between quarters on day 1, but a just significant decrement towards the end of the session as compared with performance in the first quarter of day 2 ($P = 0.05$). This hidden decrement might be attributed to 'boredom', since here the perceptual load was small. The individual trends indicating an increase in errors on day 2 were not quite significant ($P = 0.055$ with a two-tailed test), although the difference between the first and last quarters alone was significant ($P = 0.02$).

C2 only position 6 improved significantly during both the first 2 days ($P = 0.01$ or better). In both cases performance in the first quarter of the subsequent day was not

significantly better than was to have been expected from the individual trends of the previous day ($P = 0.22$ between days 1 and 2. The twenty-four subjects who carried on to day 3 performed on average in the first quarter exactly as predicted from their individual trends during day 2.) Thus there was no evidence of a decrement, either visible or hidden, during the first 2 days. If our initial hypothesis is correct (see Introduction (i)) no decrement was to have been expected in this condition, at least on the first day. For the perceptual load was of intermediate magnitude, and thus might not produce either 'mental fatigue' or 'boredom' during the $\frac{1}{2}$ hr. of the first experimental period.

There was no significant trend during day 3. Performance in the first quarter of day 4 was not quite significantly better than that to be expected from the individual trends during day 3 ($P = 0.055$ with a two-tailed test). The visible decrement on day 4 was just significant ($P = 0.05$). Only twelve out of the twenty-four subjects were still making errors, but ten of them showed the trend.

C4 only new 6's in position 6 did not yield enough data to give trends which were significant. For what it is worth, the trends shown in Fig. 1 follow most nearly those for *C3 only 6's*.

(iii) Nature of errors

C1 all new. On the first day the proportion of errors which consisted in a failure to respond rose from 46% in the first quarter to 59% in the fourth quarter. The corresponding proportions for the second day were 41 and 45%. Errors did not tend to run in sequences longer than those to be expected by chance.

Fig. 2 gives the proportion of errors for each of the critical numerals 1 to 8, and shows the numerals guessed in error. It also gives at the bottom the ratio of non-critical to critical numerals in the series. All twelve subjects made fewer errors on the numeral 7 than on 1 ($P < 0.0005$). Since out of the twenty-eight possible comparisons between numerals we have selected the most extreme, we have to divide our P value by 28. This still leaves $P < 0.014$. A Friedman two-way analysis of variance by ranks (Siegel, 1956, p. 166) showed that there were also significant differences in the proportions of incorrect guesses ($P = 0.002$). For all subjects combined, the tau correlation by numerals between incorrect guesses and the proportion of non-critical numerals in the series was 0.29, which is not significant ($P > 0.05$).

The numeral 7 was the easiest to spot, and was also one of the less likely to be guessed incorrectly. This may have been because it was the only numeral with two syllables, and also partly because it happened to be the critical signal on a relatively high proportion of its appearances in the series. The numerals 1 and 2 were the most difficult to spot. Incorrect guesses tended to avoid the numerals 1 and 8, as well as 7.

Fig. 3 gives the proportion of errors for each serial position, and the serial positions of the numerals guessed incorrectly. All twelve subjects made fewer errors on critical numerals in serial position 7 than in position 6 ($P < 0.014$). A Friedman two-way analysis of variance by ranks showed that there were also significant differences in the proportions of incorrect guesses ($P = 0.001$).

The first and last positions in the series gave the smallest proportions of errors. Guesses tended to avoid numerals in these positions. Position 4 was the most difficult position for a new numeral. Numerals here were not guessed as often as might have been expected.

C2 only position 6. Table 4 shows the errors on new numerals in the sixth position for each of the two subconditions. The total errors have been split up first according to whether the subject appeared from his previous responses to be listening to the numeral in the correct serial position or not. In each subcondition roughly one-fifth of the errors on the first day were apparently caused by listening to the wrong serial position. The remaining errors have been further subdivided according to the nature of the erroneous response. *C2a new position 6* gave a greater number of failures to change response than *C2b all position 6*. There were correspondingly reduced numbers of failures to respond, and of incorrect selections.

When a number in the wrong serial position was selected, the number in position 5 was chosen about 60% of the time, and that in position 7 about 30%. These proportions were about the same for both subconditions, and did not alter appreciably with practice.

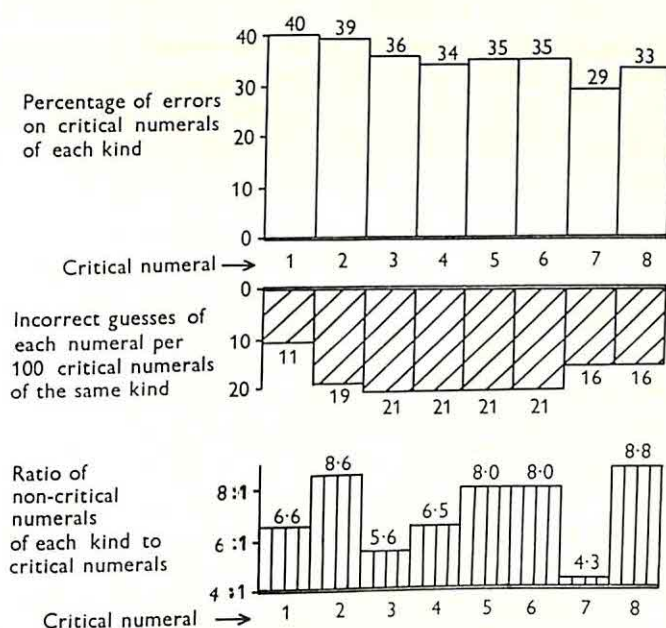


Fig. 2. The upper part shows the proportions of errors by critical numerals of each kind in C1 all new. Data from the 2 days have been combined. Means are based upon 128 responses from each of twelve subjects. The middle part shows on the same scale the proportions of guesses of each numeral, when some other numeral was correct. The bottom part shows the ratio in the experimental material of non-critical to critical numerals of each kind.

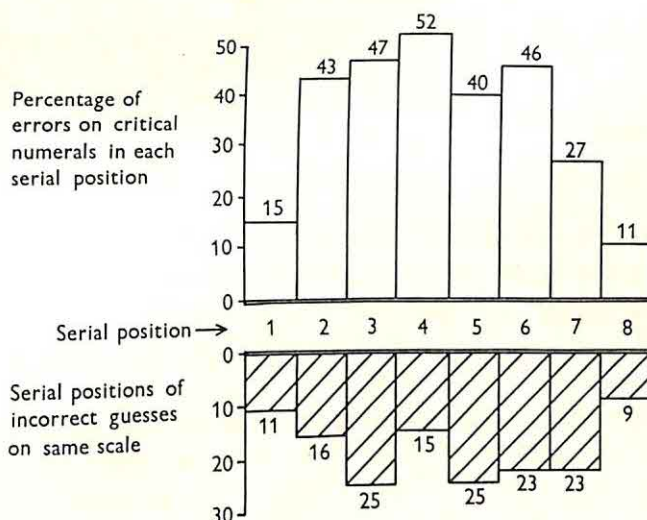


Fig. 3. The upper half shows the proportions of errors by serial position of critical numeral in C1 all new. The lower half shows on the same scale the proportions of guesses of the numeral in each serial position, when the numeral in some other position was correct.

C3 only 6's. When a new 6 was the only 6 in the series, it was missed about 2.5 times less frequently than when it occupied the same serial position, but was the second 6 in the series ($P = 0.001$). This could account for about half of the difference in this condition between scoring

all new 6's (Table 3, last column but one), and scoring only new 6's in position 6 (Table 3 last column). For new 6's in position 6 were always the only 6's in the series.

In C3b count 6's the subject had to call out after each series the number of 6's which it contained. Failures to report that a 6 had disappeared were about as frequent as failures to report a new 6.

C4 only new 6's in position 6. One-sixth of the errors were preceded fairly closely by responding in error to a new 6 in the fifth serial position. In these cases the subject may have been listening to the numeral in this serial position when the critical signal arrived. With the other errors we have no direct objective evidence to suggest what the subject was doing.

Table 4. *Nature of errors in the two subconditions of C 2 only position 6*

Classification of errors	Mean errors per 100 new numerals in position 6			
	Day 1	Day 2	Day 3	Day 4
C2a new position 6 Total	24.5	9.0	2.1	1.1
Previous response not to position 6	6.1	1.7	0	0
Previous response to position 6 Total	18.4	7.3	2.1	1.1
{ No change in response	11.3	4.8	1.4	0.9
{ Wrong numeral selected	5.2	2.0	0.7	0.1
{ No response	1.9	0.5	0	0.1
C2b all position 6 Total	26.0	9.5	5.5	2.1
Previous response not to position 6	4.6	1.1	0.3	0.3
Previous response to position 6 Total	21.4	8.4	5.2	1.8
{ No change in response	5.1*	2.7	1.2	0*
{ Wrong numeral selected	8.8	3.3	2.2	1.5
{ No response	7.5*	2.4	1.8	0.3

* C2a new position 6—C2b all position 6. $P = 0.05$ with a one-tailed test or better.

(iv) *Other points*

Scores on the practice test (see Method (iv)) correlated 0.48 ($P = 0.001$) with scores on intelligence test AH4 (Heim, 1955). The experimental material has been demonstrated to a number of foreign visitors. In order to spot the new numerals, it appears to be necessary for English to be one's native language.

IV. DISCUSSION

(i) *Decrements on day 1 were found only outside the middle range of perceptual load*

As mentioned in the Introduction (i), the hypothesis of no decrement for a task with a medium perceptual load need not apply to the second and subsequent experimental periods. It is only on day 1 that we should necessarily expect no decrement with a medium perceptual load.

C1 all new was the inspection task with the greatest perceptual load (see Table 2). The results in Table 3 support this contention. The average proportion of errors was about 40% on day 1 and 30% on day 2. This is a significantly greater proportion than in any of the other tasks investigated. Subjects performing this task for about $\frac{1}{2}$ hr. would therefore be the most likely to produce a decrement which could be attributed to 'mental fatigue'. Result (ii) shows that there were significantly more errors towards the end of day 1 than were to be expected from performance at the start of day 2. The decrement towards the end of day 1 is not visible in Fig. 1 because it was masked by learning (Thorndike, 1912).

C 3 only 6's was an inspection task with a small perceptual load (Table 2). Table 3 shows that the average proportion of errors was only about 10 % on day 1, and 5 % on day 2. This is significantly smaller than in any of the other tasks. The subjects performing this task for about $\frac{1}{2}$ hr. would therefore be the most likely to produce a decrement which could be attributed to 'boredom'. Result (ii) (see Fig. 1) shows that the trends for this task behaved very like those for C 1 all new. There was a just significant hidden decrement on day 1.

C 2 only position 6 was an inspection task with an intermediate perceptual load (Table 2). Table 3 shows that the proportion of errors on it lay between the proportions for C 1 all new and for C 3 only 6's. On day 1 it gave about 25 % of errors, and on day 2 about 10 %. As suggested in the Introduction (i), this task may be in the range of perceptual load in which we should expect no detectable deterioration during the first experimental session of $\frac{1}{2}$ hr. It may avoid both the overloading of the perceptual mechanism which should soon produce objective evidence of 'mental fatigue', and the underloading which should soon produce objective evidence of 'boredom'. Result (ii) (see Fig. 1) shows that there was no significant decrement with this task on day 1, either visible or masked by learning. This finding can be contrasted with the significant hidden decrement shown on day 1 both by C 1 all new with its greater perceptual load, and by C 3 only 6's with its smaller load.

This pattern of results is unlikely to be due to differences either in the rate or regularity of critical signals. Table 2 shows that C 2 only position 6 had a lower rate of critical signals than either of the other two tasks. Thus it was the low rate which was associated with no decrement in performance. This is the reverse of what would be predicted from previous experiments with relatively low rates of critical signals and small loads, as in C 3 only 6's (Jenkins, 1958; Kappauf & Powe, 1959; Garvey *et al.* 1959). It appears unlikely that the predicted effect of the rate of critical signals could be reversed, even for tasks with high rates of critical signals and large perceptual loads, as with C 1 all new.

The interval between critical signals was also more variable in C 2 only position 6 than in C 1 all new or C 3 only 6's (see Method (ii)). However, an increase in variability should if anything have increased the decrement in performance (Baker, 1959, Fig. 3), whereas it was only the most variable of the three tasks, C 2 only position 6, which escaped without a decrement.

The results thus support the original hypothesis of a region of optimal perceptual load. A task with a medium perceptual load did not show a detectable decrement during the first experimental session of $\frac{1}{2}$ hr., while tasks which had greater or smaller perceptual loads did show a detectable decrement. At present this conclusion has to be restricted to the particular tasks and unpractised experimental subjects used. The aim was simply to find out whether it was possible to distinguish a region of optimal perceptual load by an experiment in the laboratory. The next step would be to determine whether a similar distinction can be made within a set of closely related inspection tasks in industry.

(ii) *Motivation did not prevent lowered vigilance*

C 3 only 6's approximates a vigilance task, on account of its relatively small perceptual load and low proportion of critical signals (see Table 2). We have just seen (Discussion (i)) that performance at this task declined significantly during the first experimental period. Encouragement to the subject to try his hardest, and the

presence of the experimenter personally recording his responses, did not prevent a decrement.

C4 only 6's in position 6 was the most extreme vigilance condition investigated. It presented the smallest perceptual load, and had a very low proportion of critical signals: 1 in 64. Result (i) (see Table 3) shows that this condition was performed significantly worse than C3 only 6's, which had a greater perceptual load and a higher proportion of critical signals. If lowered vigilance could have been excluded by the experimental conditions, we should have expected the reverse result. For C4 only 6's in position 6 involved less work, and by conventional standards would thus be said to be the 'easier' task. The result found again suggests that although the experimental conditions may have encouraged the subject to try hard, this alone was not sufficient. In order to remain alert he needed a reasonable perceptual load and a higher rate of critical signals.

(iii) *Searching in the sixth serial position is comparable to searching for 6's*

Table 4 compares the two subconditions of C2 only position 6. On the first 2 days the overall proportions of errors were about the same for the two subconditions. The main differences concern the nature of the errors when the previous response was made correctly to the numeral in the sixth serial position. In C2*b* all position 6, about one-quarter of these errors involved calling the numeral which had just disappeared. In the remaining cases the subject did change his response, but he either selected the wrong numeral, or failed to respond (presumably because he had not spotted what the new numeral was).

In contrast, in C2*a* new position 6, about two-thirds of these errors involved saying 'no' (see Method (iii)) when the new numeral should have been called. There were proportionally fewer failures to respond and selections of the wrong numeral. This suggests that in this subcondition the subject may often have said 'no' when he had in fact noticed a change, because he had not spotted what the new numeral was. The data from C2*b* all position 6 thus appear to give a clearer idea of the subject's difficulties than the data from C2*a* new position 6. We shall therefore use these data alone in comparing C2 only position 6 with C3 only 6's.

Table 4 shows that the principal difficulty with C2*b* all position 6 was in spotting the new numeral. On both days 1 and 2 about 60% of all the errors involved either selecting a wrong numeral, or failing to select a numeral at all, when the previous response had been correct. The remaining errors were due either to having been listening to the wrong serial position, or to failing to note the appearance of the new numeral after having listened previously to the correct serial position. These two latter sources of error both involved failing to note the change. When combined, they give 9.7 errors per 100 responses on day 1, and 3.8 on day 2. These values are close to the errors per 100 responses in C3 only 6's: 9.6 on day 1 and 5.1 on day 2 (see Table 3). In this condition the subject had simply to call out any change (or in one subcondition any increase) in the number of 6's.

We may conclude that noting a change in the sixth serial position was of about equal difficulty to noting a change in the number of 6's. Coding by serial position and by numeral were thus more or less equivalent. The greater perceptual load of C2*b* all position 6 as compared with C3 only 6's was that the subject had to specify the new numeral in the sixth position, not merely to report that a new numeral had appeared there. The task of corresponding perceptual load involving 6's would have been to have specified the serial position of the new 6, in addition to reporting its appearance (see Table 2).

This work was carried out in the Applied Psychology Research Unit while it was under the direction of Dr N. H. Mackworth. Dr M. Stone kindly advised upon statistical methods, and also upon the exact specification of the hypotheses to be tested. The subjects were supplied by the Royal Navy. Financial support from the Medical Research Council is also gratefully acknowledged.

REFERENCES

- BAKAN, P. (1959). Extraversion-introversion and improvement in an auditory vigilance task. *Brit. J. Psychol.* **50**, 325-32.
- BAKER, C. H. (1959). Three minor studies of vigilance. *D.R.M.L. Rep.* no. 234-2. Toronto, Ontario.
- BOWEN, H. M. (1956). Unpublished Ph.D. Thesis, Cambridge University.
- CONRAD, R. (1951). Speed and load stress in a sensori-motor skill. *Brit. J. Industr. Med.* **8**, 1-7.
- DEESE, J. (1955). Some problems in the theory of vigilance. *Psychol. Rev.* **62**, 359-68.
- FISHER, R. A. & YATES, F. (1948). *Statistical Tables for Biological, Agricultural and Medical Research*. London: Oliver and Boyd.
- GARVEY, W. D., TAYLOR, F. V. & NEWLIN, E. P. (1959). The use of 'artificial signals' to enhance monitoring performance. *U.S.N.R.L. Rep.* no. 5269. Washington, D.C.
- HEIM, A. W. (1955). *Manual for the Group Test of General Intelligence AH 4*. London: National Foundation for Educational Research.
- JENKINS, H. M. (1958). The effect of signal-rate on performance in visual monitoring. *Amer. J. Psychol.* **71**, 647-61.
- JERISON, H. J. & WALLIS, R. A. (1957). Experiments on vigilance: One-clock and three-clock monitoring. *WADC. Tech. Rep.* 57-206. ASTIA Doc. no. AD118171. Dayton, Ohio.
- KAPPAUF, W. E. & POWE, W. E. (1959). Performance decrement in an audio-visual checking task. *J. Exp. Psychol.* **57**, 49-56.
- MACKWORTH, J. F. & MACKWORTH, N. H. (1956). The overlapping of signals for decisions. *Amer. J. Psychol.* **69**, 26-47.
- MACKWORTH, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quart. J. Exp. Psychol.* **1**, 6-25.
- MACKWORTH, N. H. (1950). *Researches on the Measurement of Human Performance*. London: H.M.S.O. Med. Res. Council Special Rep. Series no. 268.
- MACKWORTH, N. H. & MACKWORTH, J. F. (1958). Visual search for successive decisions. *Brit. J. Psychol.* **49**, 210-21.
- SIEGEL, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.
- THORNDIKE, E. L. (1912). The curve of work. *Psychol. Rev.* **19**, 165-94.
- WAGNER, R. C., FITTS, P. M. & NOBLE, M. E. (1954). Preliminary investigations of speed and load as dimensions of psychomotor tasks. *AFPTRC. Research Bulletin*. Tr-54-45. San Antonio, Texas.
- WYATT, S. and LANGDON, J. N. (1932). Inspection processes in industry. *Indust. Health Res. Board Rep.* no. 63. London: H.M.S.O.

(Manuscript received 15 May 1959)

ANXIETY AS A FUNCTION OF PSYCHOMOTOR AND SOCIAL BEHAVIOUR

By H. SYLVIA ANTHONY

The aim of this study was to investigate the relation between (a) psychomotor behaviour, in the form of manual response to visual signals, (b) task difficulty, and (c) social behaviour, in the form of delinquency in a military milieu. The psychomotor response was measured in terms of amplitude and duration of manual movement. Significant association was established between psychomotor and subsequent social behaviour. The delinquent groups showed overactivity, in a sense defined. Overactivity has been found typical of anxiety in other experimental and clinical observations.

The theoretical implications of the observed association are discussed in terms of Hullian learning theory and the Brown-Farber theory of frustration. It is suggested that the difference between the response patterns of the delinquent and non-delinquent groups may be accounted for in terms of the Brown-Farber formulation, under certain conditions. These require the assumption of a factor operative as a personality constant, affecting the strength of the avoidance impulse. A three-factor function of anxiety is suggested. The psychomotor and social behaviour of the delinquent group are held to show a common feature of extreme avoidance-tendency in frustrating circumstances.

I. INTRODUCTION

In an experiment which formed part of a larger investigation of occupational adaptation, Venables (1955) found that a small group of industrial employees who were subsequently discharged, largely for dishonesty, had given responses, on a psychomotor test, of a pattern significantly different from that of the total sample studied. The present research began as a replication of the experimental procedure reported by Venables, with the aim of validating this finding of significant association between psychomotor response pattern and job-unsatisfactoriness involving delinquency. The existence of such an association was confirmed in our results, as briefly reported elsewhere (Anthony, 1959). This report is mainly concerned with the theoretical formulation and implications of our findings.

In Venables's research, the finding relating to delinquency was an unexpected by-product; his experiment was designed to provide further data bearing on hypotheses largely suggested by earlier work on pilots' errors in a simulator cockpit, reported by Davis (1948). Such errors had been measured by amplitude and duration of movement, and found to range on a continuum from extensive movement in brief time to movement of long duration and relatively small extent; that is, from *overactivity* to *inertia*. An association was found between accident-proneness and extreme score at both ends of the range; also between anxiety-neurosis and overactivity, and between hysteria and inertia.

Venables aimed to find whether the overactive-inertia continuum would appear in responses made by a low-skilled industrial group to a simple, brief psychomotor task; to confirm the relation between extreme scores and anxiety/hysteria diagnoses; to find the relation between introversion-extraversion and overactivity-inertia, and the reaction of (a) neurotic patients, and (b) selected introvert and extravert industrial subgroups to changes of task-difficulty.

Performance on the task was found to fall on a continuum of overactivity-inertia,

using a score which will be more fully described later. The neurotic patients showed significant differences of score at the expected extremes of the range. The anxiety-state patients, initially overactive, became even more so when task-difficulty increased, and reverted to their initial level when it was reduced again. The hysteric patients, initially inert, remained inert when task-difficulty increased, but became more active when it was reduced at the end of the test-session. The behaviour of industrial groups selected as extreme cases on tests of neuroticism and introversion-extraversion, was somewhat ambiguous. The extravert group scored almost identically with the anxiety-state patients on the easy task, but more like the hysteric patients in manner of reaction to difficulty; the introvert group showed greater inertia, finally scoring at the low level of the hysteric patients, but in manner of reaction to difficulty they resembled the anxiety-state patients.

The behaviour of our delinquent groups resembles that of Venables' extravert sample. He named this sample *quasi-hysteric* in his report. The description is not applicable to delinquent groups in our research, since they are more significantly characterized by overactivity (in which respect they resemble anxiety-state patients) than by manner of reaction to change of difficulty of task (in which respect they resemble hysteric patients). This fact, and the ambiguity of Venables's findings at this point, leads us to reject hypotheses explaining association between delinquency and psychomotor behaviour in terms of Eysenckian dimensions of introversion-extraversion or neuroticism, and to seek a theoretical formulation of the observations in other terms.

Davis (1948, p. 39) had attributed pilot error to 'anticipatory tension, closely related to anxiety'. The differences observed between our delinquent and non-delinquent groups corresponded in many ways to those reported between anxious and non-anxious subjects in experimental studies carried out within the framework of Hullian theory. An interpretation was therefore sought mainly in terms of that theory.

II. EXPERIMENTAL PROCEDURE

The apparatus of the psychomotor test used in this investigation has been previously described by Venables (1955), and since the same apparatus was used and the procedure replicated as closely as possible, some details will be omitted from the present description.

Administration is individual. The actual test takes under 5 min. for each subject, with a preceding practice-period of 3 min.; in the special trade-selection programme of which it formed a part, 15 min. total time was allotted, so that there should be no sense of rush, or lack of time for establishment of rapport between the candidate and the junior N.C.O. administering the test. The special selection procedure did not include other tests; the subjects usually came to the psychomotor-test session from a waiting-room where papers and games were provided and talking permitted with other candidates.

Apparatus. From the seated subject's viewpoint, this appears as a large box placed before him on a table. At about elbow-level is a shelf in which a slot $8\frac{1}{2}$ in. long runs left to right; from this slot protrudes a movable knob with a pointer attached. Behind the shelf, a framed milk-white screen $8\frac{3}{4}$ in. wide rises at an angle of 45° . On the base of its frame are three markers, one central and the others equidistant $2\frac{3}{4}$ in. from it. Thus outwards from the marker-positions, the slot and the screen continue for approximately 2 in., and it is possible for the pointer to overshoot the mark by this amount.

Behind the screen are placed, over each of the side markers, vertical columns of five stimulus-lights, in form rather like miniature traffic-signals. Their position is not visible until alight. Reading from the bottom the lights show (1) white, (2) green, (3) red, (4) YES, (5) NO.

Instructions are as follows. 'When you hear this click' (a signal indicating that a new stimulus is being presented) 'move the pointer from the mark in the centre to the mark under the white light, if it is on its own, or has green or YES above it. But if it has red or NO above it, move the pointer to the mark on the opposite side. Then move the pointer as quickly as possible back to centre.' The instructions are repeated, with a final reminder: 'There and back as quickly as possible.' A period of practice ensues, at slow speed (3 sec. intervals between signals), during which the subject may ask questions. He is then told that there will be a series of white-light signals, followed by a slightly faster series using all the lights, and then a third series using white lights only as before.

The signals are given in a predetermined random sequence, fifty presentations in each of the three series, *A*, *B* and *C*. The first and last series are identical sequences given at 2 sec. intervals. *B*-series signals are given at 1.5 sec. intervals, and employ all the lights except on every fifth display, when white lights only are used, as in *A* and *C* series. Only responses to white lights are scored, the response to every fifth signal being scored throughout the test.

The pointer is attached to a lever controlling a pen which records the subject's response on a moving paper. Another pen, mechanically controlled, traces a central line. The trace made by the subject's handling of the knob between responses also runs approximately centrally along the paper. The response-movements trace hairpin-shaped deviations from this central line, their extent being due to the amplitude of the movement of the hand, and their basal width to the movement of the paper, registering the time spent between the beginning and completion of the response.

The score for each of the three series is obtained by: (a) measuring, in every fifth response-trace, the sideways-extent and the basal width; (b) summing the ten measures so obtained in each case; (c) calculating the ratio of the sums, extent/width; and (d) taking the log to base *e* of this ratio (Fisher & Yates, 1957, table XXVI)—to two decimal places.

The use of this ratio as test score is due to Venables, who experimented with several combinations of the two measures, which he symbolized as *R* (response) and *W* (width). We logged the ratio so as to normalize score distributions otherwise consistently skew, and symbolized the components of the logged ratio as *E* and *T*, so as to avoid confusion with the unlogged measures, and to record the view that the measure of width is a measure of time. The variates E/T_A , E/T_B , E/T_C , multiplied by 100, will be referred to as *A*, *B* and *C* scores.

The subjects were young men presented to a Trade Selection Board as candidates for recruitment to a Royal Air Force ground trade, at a period when almost all men aged between 18 and 25 years were liable for two years' military service. The modal age on testing was between the eighteenth and nineteenth birthday; the upper limit of range was 34, but few candidates were over 23 years of age. Conduct-records of samples entering within a given period were finally analysed when those entering on the minimum 2-year engagement were due for discharge, so as to maintain the character of the sample in respect of engagement undertaken.

Every candidate had been recently passed fit by Medical Boards, and accepted as eligible for R.A.F. ground-trade service on the results of recruit-tests of ability and aptitude. Candidature was voluntary; those not genuinely desirous of entering the trade were not favourably considered by the Board.

The trade demanded certain qualities of character or temperament rather than special technical skill or educational attainments. Recruits of exceptional intelligence or skill were mostly earmarked for other trades or for officer-selection. In consequence, the sample, while excluding the dullest recruits, included a relatively high proportion of men occupationally unsettled or poorly qualified. It also included, however, a number whose skills fitted them for civilian trades or professional training in which the R.A.F. could not currently place them, and a few who had junior civil experience in the trade for which selection was being made. The average general intelligence of candidates, as measured by R.A.F. entry tests, approximated that of the United Kingdom population of their age-group, but with relatively little variance, few scoring below -0.5 or above $+1.0$ z-score.

Over a period of more than two years, more than 2000 men were tested. Records of subsequent conduct were available for those who entered the trade—over 75% of the total. The samples quoted were drawn from the body of candidates either by random sampling towards the close of the period, or by using all records of candidates or entrants tested up to a given date. In addition to the candidates, 20 subjects were tested as controls; unfortunately the opportunity for testing non-candidates was very limited, after the apparatus had been checked and training given on administrative procedures.

The criterion of delinquency. The criterion of delinquency was dual: by number of occasions of offence (convictions), and by penalty awarded, as shown in the list below.

Category	Convictions	Penalty
(o)	0	None
(a)	1	Admonished or reprimanded only
(b)	1	More than (a), less than (c)
(c)	1	Confined to camp 7 days; loss of 1 day's pay or both penalties
(d)	1	More severe than one of (c)
(e)	2 or more	All less than (c)
(f)	2	(c) or (d) on at least one occasion
(g)	3 or more	Same as (f)

Penalties above category (c) may be awarded only by senior officers or civil courts; such offences are listed in Table 1 as serious. Minor offences include behaviour which in civil life would rarely be classed as crime, though if persistent it might result in job-dismissal or unemployment. Those judging offenders knew nothing about the test.

Table 1. *Number of convictions and nature and rating of offences of candidates who entered the trade over a period of 5 months*

Nature of one offence per conviction	Rating of offence			
	Serious		Minor	
	No.	%	No.	%
Theft, arson, desertion, striking a superior	5	13.9	—	—
No information	2	5.6	—	—
Absent without leave, more than 24 hr.	12	33.3	—	—
Absent without leave, less than 24 hr.	2	5.5	41	28.9
Neglect of duty, disobeying orders, conduct prejudicial to discipline, etc.	8	22.2	70	49.3
Motor transport (including civil)	6	16.7	7	4.9
Civil (not M.T.)	1	2.8	2	1.4
Dress, kit, loss of documents, etc.	—	—	22	15.5
Total (convictions)	36	100.0	142	100.0
Total (offenders)	34		83	

Notes. (1) The offenders totalled 117, representing 28 % of all candidates entering the trade over the 5-month period, the conduct records of the total sample having been studied over a period of approximately 1½ years after trade-entry. The remainder of the sample had no record of conviction. (2) Only one of several offences simultaneously, the penalty may be more severe than would normally be awarded for a single offence. This method prevents exactitude in the numerical distribution by nature of offences.

III. EXPERIMENTAL RESULTS

Preliminary findings. Distributions of E , T , and E/T scores were found to be continuous and unimodal, tending to bunch at the lower end of the scale in each case, but approximating symmetrical form when logged.

The score variate E/T , being an average of distance moved over time taken, is obviously greatest when an extensive movement is made in brief time, and least when a movement falling short of the mark takes a long time to complete. Since the subject is instructed to make a movement of precise extent as quickly as possible, optimum performance is represented by central E values divided by minimum T values; the same score would be given by a larger E divided by a larger T . A wide

range of scores will clearly have the same value as that produced by optimum performance; in effect, the scale does not measure performance efficiency. Further, E and T are measured and scored irrespective of whether the subject moves the knob in the right or wrong direction.

From the physical properties of the apparatus manipulated, one would expect that, as a general rule, more time would be taken in making a movement which overshoots the mark than in making one which falls short of it, and vice versa. A positive correlation would then appear between E and T values, and the ratio E/T would tend to minimize the differences due to E and T considered separately. This does not occur. In a sample of 169 candidates tested to a given date, the product-moment r_{ET} was found to be negative and significant at or above 1% level for both A and B series. Subjects who tend to overshoot, tend to take less time, and vice versa. The ratio, by combining the two measures, therefore provides an index more extensive than either measure taken separately.

Venables's finding of non-significant correlation between psychomotor test scores and general intelligence was confirmed in respect of A , B , and C scores and reaction to change of task-difficulty ($A-B$, $B-C$).

Individuals tend to score at a characteristic level throughout the test. In a random sample of 20% of all candidates tested to a given date (sample $n = 251$), the product-moment r coefficients of series' scores were: AB , +0.88; AC , +0.87; BC , +0.87.

In previous researches, the direction of deviation of B score from A and C scores has been found to be associated with personality factors: the confirmation of this finding is shown below.

Association between test results and conduct records. Study 1. A sample of all candidates who entered the trade for training up to a given date ($n = 601$), was divided into *good* and *bad* groups. The former included criterion categories listed (o) to (c) above, together with a small group who had failed on trade training for technical reasons only. The latter included the remaining criterion categories, together with a small group who had failed on trade training owing to unsuitable character as judged by the training staff. The staff had no knowledge of the test results.

Differences of mean score between categories *within* each group were not significant, in the sense that the value of P was in every case greater than 5%. Differences of mean score between the *good* and *bad* groups were statistically significant, as shown in Table 2 below.

The *bad* group is found to make a significantly more active score on the initial series (A); to reduce activity when the task becomes more difficult on B series,

Table 2. Mean and variance of scores of 'good' and 'bad' groups (Study 1)

Series	'Good' group		'Bad' group		P of difference between means
	Mean	Variance	Mean	Variance	
A	196	(1494)	220	(1604)	< 0.001
$A-B$	3.2	(420)	13.0	(629)	< 0.05
$B-C$	-6.2	(439)	-14.0	(347)	< 0.05
n	567		34		

significantly more than the *good* group; and to revert to approximately the former activity level when the task becomes easier again and the end of the session is in sight.

Despite the relatively large swing towards inertia on the difficult task, the *bad* group made a mean score significantly higher than that of the *good* group on the difficult task also ($P = 0.05$).

Later studies. The general picture presented in Table 2 was maintained as greater numbers passed before the Board, and the trade-entrants (whose records alone could be followed up) moved into and sometimes on from the successive crime categories. Comparison of the no-conviction and one-minor-crime groups consistently indicated no significant score difference between them; the former group has therefore been omitted from the study quoted in Table 3. The training-withdrawal groups have also been excluded from the sample, owing to the lack of further records in their case. Scores of a small group of non-candidates have been added.

Table 3. *Mean and variance of scores of groups distinguished by crime-category and candidature (Study 2)*

Series	Non-candidates		Candidates						
			<i>abc</i>		<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>defg</i>
	$\bar{x}$	s^2	$\bar{x}$	s^2	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$	$\bar{x}$ s^2
<i>A</i>	177*	(894)	194	(1789)	203	215	227	251	217*** (1331)
<i>B</i>	187	(1196)	192	(1422)	198	209	215	222	208*** (748)
<i>C</i>	182	(1391)	197	(1781)	210	218	219	244	219*** (879)
<i>A-B</i>	-10.0***	(233)	+2.1	(228)	+7.1	+1.5	+10.6	+28.5	+9.0 (600)
<i>B-C</i>	+6.0**	(353)	-5.8	(318)	-11.1	-9.5	-2.7	-22.0	-10.0 (375)
<i>n</i>	20		93		17	10	9	4	40

Note. The difference from the corresponding score of the *abc* (one-minor-crime) group is significant at * 5 % level, ** 2 % level, *** 1 % level.

The difference between the scores of non-candidate and *defg* (serious and multiple crime) groups is in every case significant at or below 5 % level. (Fisher & Yates, 1957, Table III.)

It is seen that the non-candidate group showed lower activity scores, with increase of activity when task-difficulty increased. The differences between delinquent and non-delinquent (one-minor-crime) tradesmen confirm in more detail those already shown by the results in Table 2.

Differences in score-variance between groups were rarely found significant at 5 % level; the number of instances of difference approximating this level were not of consistent direction, and appeared attributable to chance.

IV. DISCUSSION

The score as a measure of anxiety. The psychomotor test score is a measure of activity, defined in terms of amplitude and duration of movement. In the interpretation of researches on similar tests Davis (1946, 1948) used the concept of anxiety, referring particularly to the work of Mowrer (1939) on the relation between heightened expectancy and overactivity. Anxiety is here defined as a hypothetical state accompanying or resulting from the anticipation of noxious stimulation such as pain, restriction of functional freedom, diminution of status or failure to attain a goal.

The association between conditions producing such anticipation, and muscular tension, restlessness and lowered stimulus-thresholds has been frequently demonstrated as, for example, by Freeman (1939) and Masserman (1943), and more recent experiments have shown the association between anxiety manifested verbally and rapidity and amplitude of manual movement (Castaneda, 1956). The relationship of anxiety and activity does not appear to be simple and constant; our results confirm those of Davis in showing that high initial activity, apparently associated with anxiety, is under certain conditions followed by extreme reduction.

Anxiety as drive. In Hullian learning-theory, anxiety is frequently identified with *drive*, and is held to take the place, as motivation for human action, of the primary biological urges such as hunger. As such, it represents primarily a genetic, unlearned component of behaviour. From the interpretation of *drive* as anxiety, it is predicted that more-anxious individuals, or individuals when more anxious, will learn more rapidly or efficiently than less-anxious individuals, or the same individuals when less anxious. In the comparison of groups of subjects whose level of anxiety has been determined either on a verbal scale or by other methods, this prediction has frequently been verified with regard to speed of conditioning and the learning of simple tasks, but when more complex demands are made on the subject, the reverse has often been found; subjects judged less anxious tend to give superior performance. From the extensive literature may be mentioned the studies of Welch & Kubis (1947), Lucas (1952), Montague (1953), Farber & Spence (1953), Spence & Taylor (1953), Taylor (1956), Davidson, Andrews & Ross (1956), Hill (1957), and Nicholson (1958).

The psychomotor test does not measure current learning, nor is the score an index of efficiency; nevertheless, if the initial activity score is taken as an index of chronic anxiety level, a corresponding pattern appears to that presented by the results in researches on the relationship between anxiety and learning proficiency. The groups scoring higher on the simpler task tend to reduce their activity on the succeeding complex task, while the group scoring lower on the simpler task tend to maintain or increase their activity. The pattern appears in Table 3 above, if delinquency is taken as the criterion of anxiety.

Anxiety as stimulus to habitual response. The reduction of score of the more-anxious groups on the more difficult task is not easy to reconcile with a theory identifying anxiety with drive alone. Montague (1953), faced with the apparent anomaly of relatively low scores of more-anxious subjects on a relatively complex task, suggested that stimuli associated with anxiety itself may elicit specific responses—visceral, muscular, etc. Lucas (1952) suggested that ‘anxiety stimulates responses of a predominantly aggressive nature, or avoidance-responses learned from the consequences of aggression’. Nicholson (1958) concluded from experimental evidence that anxiety operated under certain conditions as stimulus to responses which interfered with learning efficiency, and that such responses, aroused by anticipation of failure, were more likely to occur when the task was difficult, and when instructions favoured ego-orientation.

The conditions in which (by these hypotheses) anxiety may stimulate specific habitual responses, and the type of response then elicited, may occur in the psychomotor test situation. The effect of anxiety on learning-proficiency, in relation to task-difficulty, is similar to its effect upon psychomotor behaviour. The reduction of

activity here observed as characteristic of more-anxious subjects, may be an early stage in a process which, if repeated, would tend to place the more-anxious below the less-anxious on any register of efficiency of responses which may be promoted by psychomotor activity.

Anxiety in relation to frustration. As voluntary candidates for recruitment to the trade, subjects set themselves the goal of demonstrating their capacity. In the test, which is a part of the special selection procedure, the subject is required, on receipt of appropriate cues, to move his hand (a) quickly, (b) with spatial precision, and (c) correctly left or right. In the second task-series, the cues by which the correct direction of movement is to be determined are in most instances more complex than in the first series. Under such conditions, reaction-time between perception of the signal and inception of the responsive movement is normally longer (Berlyne, 1957). There is then less time remaining between successive signals, for the completion of the required response. At this stage of the test, when the subject requires more time in order to comply efficiently with instructions, less time is given; the period between signals is cut down by a quarter. The subject is frustrated in his effort to comply with instructions and to demonstrate his capacity. This occurs in a series of fifty successive signals, of 75 sec. total duration, during which anticipation of failure may be roused.

When the signals are thus speeded up, the least-anxious group ($n = 20$ in Table 3) tend to speed up their responses, giving a higher mean score on *B* than on *A* series. The more-anxious group ($n = 40$ in Table 3) tend to reduce speed, giving a lower mean score on *B* than on *A* series. They do not respond more slowly on *B* than the less-anxious, nor is their initial score so high that it cannot be maintained or exceeded—or so one would deduce from Venables's (1955) observations of the behaviour of anxiety-state patients. But while the less-anxious tend to change their activity in accordance with the change in the cue for action, and so compensate for the reduction in available time, the more-anxious do the reverse. The former show positive adaptation, the latter not.

When frustration is due to sanctions imposed by established authority, correspondingly non-adaptive behaviour is liable to issue in delinquency. Neither the psychomotor nor the social behaviour of the delinquent group appears as substitutive or symbolic; both appear to be the expression of a constant manner of response to frustration, differing in a specific way from the response of the major group. The theoretical significance of the difference may be studied by fitting the test results to a frustration-model such as that proposed by Brown & Farber (1951).

According to this model, the response to a frustrating situation may be considered as the resultant of conflicting impulses of *approach* and *avoidance*, the stronger of which determines the direction of action. In social situations, in which aggressive actions of approach may be a way of avoiding compliance with sanctions, it is necessary to define *approach* further, as adaptation to, or acceptance of the social demand, and *avoidance* as non-adaptation to it.

In relatively simple conditions, avoidance appears as a reaction to anxiety about the consequence of pursuing an initially stronger approach-impulse. In social situations, approach may be due to anxiety about the consequences of avoidance. In general, it may be supposed that conflict between impulses is associated with anxiety proportionate to its severity.

The strength of the *sense of frustration* is given by the Brown-Farber formula:

$$F = \frac{E_w^n}{E_s^{n-k}},$$

where E_w represents the excitation aroused by the weaker and E_s that aroused by the stronger of two competing impulses; in illustrating the formula the authors evaluate n as 2, k as 1. Following Hullian learning-theory, excitation is held to be in each case the product of appropriate drive and of stimulus to habitual responses aroused by the immediate situation.

As the major difference observed between group scores occurs on A series, while the main objective frustration occurs on B series, the differences obviously cannot be completely accounted for by reaction to objective frustration. They may, however, be explained as a product of F and a constant personality factor which tends to increase the strength of the weaker impulse, and has higher values in delinquent subjects. Let us make three assumptions: (a) that the approach-tendency predominates in all subjects on entering the test (approach = E_s); (b) that objective frustration tends to increase the strength of the weaker impulse; (c) that owing to the previous operation of the constant factor, the approach and avoidance tendencies are of more nearly equal strength in the delinquent than the non-delinquent group.

It follows from these assumptions that the avoidance tendency may become stronger than the approach tendency in response to frustration during the test, and that this is more likely to occur in the delinquent than in the non-delinquent group.

By definition of F , it follows that at any given level of total excitation, the sense of frustration derived from impulses of approximating strength is greater than that derived from impulses of which one clearly predominates. From the assumptions above, it then follows that the value of F will be higher in the delinquent than in the non-delinquent group on entering the test. The constant factor is also assumed to have higher values in the delinquent group; A score, as product of F and the constant, will therefore also tend to be higher. This accords with observation. When frustration occurs, on B series, the strength of the avoidance tendency is assumed to increase in both groups, and may then become predominant in the delinquent group, in which the difference of strength of the two tendencies was previously relatively small. If this occurs, the positions of approach and avoidance tendencies in the F ratio are reversed, and the direction of behaviour falls to be determined by the avoidance impulse. If the avoidance tendency becomes stronger than the approach tendency previously was, the value of F , and hence the value of the activity score, will be reduced. These are the qualitative and quantitative forms of response actually observed in the delinquent group. At the other extreme, in the least-anxious group (non-candidates in Table 3) in which a relatively weak avoidance tendency is to be assumed, its increase with frustration will theoretically lead to increase in value of F and of activity score, as actually observed.

By this theory, repetition of frustration would be expected to produce reduction of score also in less-anxious subjects, similar to that observed here in the more-anxious. This appears to have occurred in Davis's more protracted experimental situations, where the main difference lay in *when*, not *whether*, inertia supervened.

As the product of F and a factor characteristically strong in more-anxious subjects,

the activity score appears to represent a three-factor anxiety function. That the observations may be represented by such a function is graphically illustrated in Fig. 1, where scores of groups shown in Tables 2 and 3 above are entered according to the formula $Y = x^3/t$, in which Y represents psychomotor test-score in standard deviation units; t represents task-difficulty, the more difficult task having the higher value; x represents hypothetical anxiety-level in deviation units.

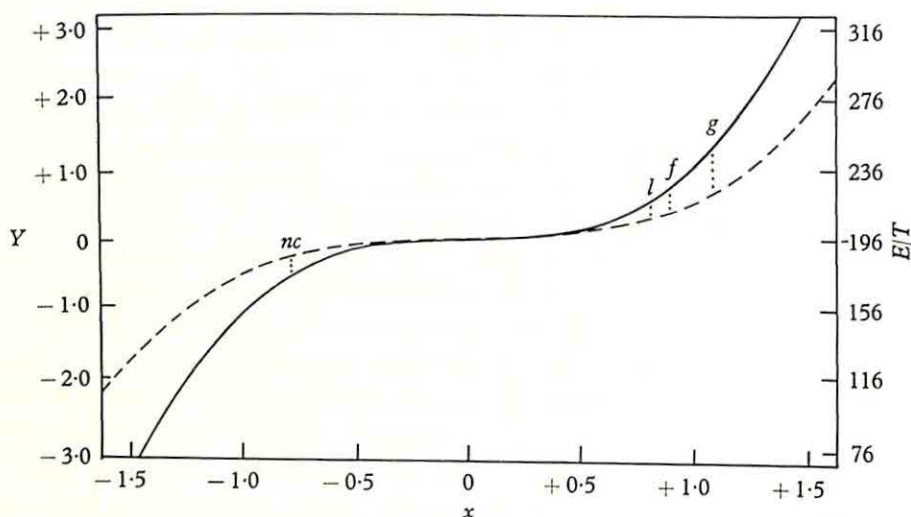


Fig. 1. The curves show Y and E/T score on A and B test series in relation to score on the x -variate. 'Good' candidates' mean scores (not shown) approximate zero on both co-ordinates. —, A score ($t=1$); ---, B score ($t=2$); . . . , group mean scores. Groups: nc , non-candidates (Table 3); l , 'bad' group (Table 2); f , g , groups f and g (Table 3).

The formula offers an extreme over-simplification of the relations of the factors assumed to be operating to produce the test-score.

The object of anxiety. The object of the test is to assess personal capacity. This, or negatively, ego-impotence, is also assumed to be the object of the anxiety displayed in the psychomotor and the delinquent behaviour.

V. SUMMARY AND CONCLUSIONS

Delinquent groups have been found to differ significantly from non-delinquent in manner of response to a psychomotor test. On an easy task the delinquent group shows greater activity, in a sense defined. Significant reduction of activity follows when the task becomes more difficult, and reversion to approximately the former activity-level when it becomes easier again.

Overactivity has been shown, in other researches, to be a characteristic form of expression of anxiety. The delinquent group, whose test-scores indicate overactivity, may therefore be classed as more-anxious.

The reduction of activity of the more-anxious group on the difficult task corresponds to the relatively poor performance reported of more-anxious subjects on complex learning-tasks; in both cases anxiety appears to function as both *drive* and *stimulus to responses habitually aroused by fear of failure*. Reduction of psychomotor activity may be a condition of the poorer learning performance found in more-anxious subjects on complex tasks.

Analysis of the current experimental responses indicates that the least-anxious group act in a positively adaptive way when faced with increased difficulty, while the delinquent group do not. In this respect the psychomotor behaviour of the delinquent group in reaction to frustration resembles its social record.

The differences of reaction of the delinquent and non-delinquent groups on the easy and difficult tasks, as represented by their mean scores, may be accounted for as products of (a) sense of frustration (F) and (b) a constant personality-factor. The former may be defined by the Brown-Farber formula as a ratio of competing approach and avoidance tendencies, and the latter as chronic anxiety, tending to increase the initially weaker avoidance-impulse. A three-factor anxiety-function is indicated. By the same theory may be explained Davis's observations (1948, 1958) of inertia as a phase tending to occur in both more and less-anxious subjects when frustration is maintained over a considerable period.

From its occurrence in a situation designed to test personal capacity, the constant anxiety factor is assumed to be, in extreme form, fear of ego-impotence, or acute lack of self-confidence. Such fear, if conditioned by generalized early experience rather than by frank objective inferiority, would normally be extinguished by maturation. The pre-selection of candidates by Medical Boards minimizes objective differences in relevant capacity. Persistence in the young adult of anxiety about ego-potency may be due to development of a habit of avoiding situations in which personal capacity is objectively tested, in consequence of which he does not become conscious of the maturation of his powers. Hypothetically such a habit may be attributed to exceptionally discouraging early experience. The question whether, in our experimental population, the reduction of score, interpreted as avoidance-behaviour, is associated with adverse early experience, remains as a question for further research.

The psychomotor and social behaviour of the delinquent group are linked by the common feature of unusually strong avoidance tendency in frustrating circumstances, which in the social situation takes the form of absence, negligence, or other disregard of sanctions, and in the test, the form described. This behaviour of the delinquent group is emphasized and illuminated by the contrast with its typical extremely active, apparently co-operative behaviour in the less exacting task. The normal (non-delinquent) subjects tend to present a lower general psychomotor activity-level, with positive adaptation to frustrating circumstances in both the test and the social situation.

The writer gratefully acknowledges the assistance of Dr P. H. Venables (who lent the psychomotor apparatus), and the invaluable skilled co-operation of R.A.F. personnel.

REFERENCES

- ANTHONY, H. S. (1959). Association between psychomotor behaviour and delinquency. *Nature, Lond.*, **183**, 343-4.
- BERLYNE, D. E. (1957). Conflict and choice time. *Brit. J. Psychol.* **48**, 106-17.
- BROWN, J. S. & FARBER, I. E. (1951). Emotions conceptualized as intervening variables—with suggestions towards a theory of frustration. *Psychol. Bull.* **48**, 465-95.
- CASTANEDA, A. (1956). Reaction-time and response-amplitude as a function of anxiety and stimulus-intensity. *J. abnor. Psychol.* **53**, 225-8.

- DAVIDSON, W. Z., ANDREWS, T. G. & ROSS, S. (1956). Effects of stress and anxiety on high-speed color-naming. *J. exp. Psychol.* **52**, 13-17.
- DAVIS, D. R. (1946). The disorganization of behaviour in fatigue. *J. Neurol. Psychiat.* **9**, 23-8.
- DAVIS, D. R. (1948). *Pilot Error*. H.M.S.O.
- DAVIS, D. R. (1958). Clinical problems and experimental researches. *Brit. J. Med. Psychol.*, **31**, 74-82.
- FARBER, I. E. & SPENCE, K. W. (1953). Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.* **45**, 120-5.
- FISHER, R. A. & YATES, F. (1957). *Statistical Tables for Biological, Agricultural and Medical Research*, 5th ed. Edinburgh: Oliver and Boyd.
- FREEMAN, G. L. (1939). Postural tension and conflict situation. *Psychol. Rev.* **46**, 226-40.
- HILL, W. F. (1957). Comments on Taylor's 'Drive theory and manifest anxiety'. *Psychol. Bull.* **54**, 490-3.
- LUCAS, J. D. (1952). The interactive effects of anxiety, failure and intra-serial duplication. *Amer. J. Psychol.* **65**, 59-66.
- MASSERMAN, J. H. (1943). *Behaviour and Neurosis*. University of Chicago Press.
- MONTAGUE, E. K. (1953). The role of anxiety in serial rote learning. *J. exp. Psychol.* **45**, 91-6.
- MOWRER, O. H. (1939). A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.* **46**, 553-65.
- NICHOLSON, W. M. (1958). The influence of anxiety upon learning: interference or drive increment? *J. Personal.* **26**, 303-19.
- SPENCE, K. W. & TAYLOR, J. A. (1953). The relation of conditioned response strength to anxiety in normal, neurotic and psychotic subjects. *J. exp. Psychol.* **45**, 265-72.
- TAYLOR, J. A. (1956). Drive theory and manifest anxiety. *Psychol. Bull.* **53**, 303-20.
- VENABLES, P. H. (1955). Changes in motor response with increase and decrease in task-difficulty in normal industrial and psychiatric patient subjects. *Brit. J. Psychol.* **46**, 101-10.
- WELCH, L. & KUBIS, J. (1947). Conditioned PGR in states of pathological anxiety. *J. nerv. ment. Disease*, **105**, 372-81.

(Manuscript received 18 February 1959)

NEUROTICISM IN CHILDREN

II. IDENTIFICATION AND MEASUREMENT OF A NEUROTICISM FACTOR*

By J. G. THORPE

Banstead Hospital, Sutton, Surrey

In the first of these papers dealing with neuroticism in children (Thorpe & James, 1957) it was shown that the scores made by children on the Maudsley neurotic test battery were to a large extent situationally determined. Two groups of children attending the Maudsley Child Guidance Clinic were given a selection of these tests, one group was tested away from the hospital, and the other was tested within the hospital. The results indicated that the testing situation affected the responses of neurotic children in such a way as to make them appear more neurotic than a similar group tested outside the hospital situation. Not only were the questionnaire responses affected in this manner, but also the responses to the objective performance tests.

Since in all previous studies neurotic children had generally been tested in places associated with their emotional disturbance, and normal controls in places apparently lacking this association, it was obviously of importance to determine whether the differences in the test scores obtained by these normal and neurotic children could be accounted for in terms of the effects of the testing situation alone. Two groups of children, one neurotic, the other normal, were therefore tested in their own schools. It was not disclosed to these children that the examiners were in any way connected with the Maudsley Hospital. The conclusions of the investigation were that when the influence of the testing situation upon test scores was taken into account the questionnaire type of test was of no value in discriminating between normal and neurotic children, though this conclusion did not apply to the objective performance tests, most of them showing the sort of group differences which were expected. It should be noted, however, that the obtained differences were smaller than had previously been the case.

In the present paper the main task was to continue the analysis of the test data presented in the first paper. As previous authors have failed to control the effects of the testing situation, their resulting 'factors' of neuroticism, whether obtained by orthodox Factorial analysis (Cox, 1951), by Canonical Variate analysis (Himmelweit & Petrie, 1951) or by Criterion analysis (Connor, 1952) must necessarily have been influenced by these uncontrolled situational factors. Our present data, uncontaminated by these influences, can therefore be expected to produce valuable and important results.

The test scores of the 34 neurotic children and the 106 controls were intercorrelated and age was partialled out. This matrix was then factored by the centroid method

* This work was carried out at the Maudsley Hospital under the direction of Prof. H. J. Eysenck and financed by the Central Research Fund of the Bethlem Royal and Maudsley Hospitals.

and five significant factors were extracted. Both Tucker's and McNemar's criteria indicated that further factors would have proved statistically non-significant. A consideration of the centroid solution failed to suggest that any one of these factors could reasonably be labelled 'neuroticism'. That is, the tests of neuroticism did not show higher loadings on any one of the factors than they did on the others.

As the total battery comprised neuroticism, intelligence, and psychoticism tests, a subsequent analysis into simple structure components was attempted in the hope that these dimensions would become apparent. Thurstone's method of extended vectors was used in conjunction with a Landahl rotation, but the analysis had to be abandoned when it became clear that no *unique* simple structure solution to the matrix could be found.

At this juncture use was made of the column of critical ratios given in our first paper by resorting to Criterion analysis (Eysenck, 1950). Embodied in the general principle of Criterion analysis is the idea of rotating a factor into maximum correlation with a criterion, i.e. rotation in such a way that the discrepancies between the factor saturations and the power of each test to discriminate between the criterion groups is minimized. The extent to which this minimization is effective is usually measured by correlating the column of critical ratios with the column of rotated factor saturations, and the value of this correlation will determine the amount of conviction we have in giving to the factor the name that we gave to the criterion.

The centroid solution was accordingly rotated in this manner. The saturations on the rotated factor correlated 0.492 with the criterion column. With an N of 32 a correlation of 0.45 is required for significance at the 1% level. Our obtained correlation is therefore significant on this test at beyond the 1% level* and indicates that the rotated factor can reasonably be labelled 'neuroticism'.

Factor scores on the neuroticism factor were then estimated for all children in the two groups. Only tests having loadings significant at or beyond the 5% level were used in the factor estimation, and equal weights were employed. These factor scores suggested that there might be a real difference between the two groups, though the 5% level of significance was not reached.

The conclusions emerging from this analysis are as follows. (i) When the testing situation is carefully controlled the factor of neuroticism is much more difficult to identify. (ii) Although the present results are suggestive of the existence of a dimension of neuroticism, even when the effects of the testing situation have been controlled, they also suggest that much confirmatory work is required. (iii) Until this work has been carried out, the satisfactory measurement of the factor will be bound to prove elusive.

* The final correlation between factor and criterion might more properly be regarded as a *multiple* correlation. In the analysis the correlation is computed between each column of factor saturations and the criterion column, and each factor is weighted roughly according to the size of this correlation. When the final correlation between factor and criterion is so regarded, and the standard error for the multiple coefficient is used in the evaluation of this correlation, an R of 0.492 will be just significant at the 5% level.

REFERENCES

- CONNOR, D. V. (1952). The effect of temperamental traits upon the group intelligence test performance of children. Ph.D. Thesis, University of London Library.
- COX, S. (1951). A factorial study of the Rorschach responses of normal and maladjusted boys. *J. genet. Psychol.* **79**, 95-115.
- EYSENCK, H. J. (1950). Criterion Analysis—an application of the hypothetico-deductive method to factor analysis. *Psychol. Rev.* **57**, 38-53.
- EYSENCK, H. J. (1952). *The Scientific Study of Personality*. London: Routledge and Kegan Paul.
- HIMMELWEIT, H. & PETRIE, A. (1951). The measurement of personality in children. *Brit. J. educ. Psychol.* **21**, 9-29.
- THORPE, J. G. & JAMES, D. P. (1957). Neuroticism in Children. I. An investigation of normal and neurotic group differences. *Brit. J. Psychol.* **48**, 26-34.

(Manuscript received 25 November 1958)

DELINQUENCY, MALADJUSTMENT AND UNFAVOURABLE ECOLOGY

By D. H. STOTT

Psychology Department, Glasgow University

The dichotomy of 'psychological' and 'sociological' delinquency is criticized. To ascertain to what extent delinquency is purely cultural, the prevalence of psychological factors was investigated in a year-group of Glasgow boy probationers. The latter were found to be strikingly more prone to behaviour-disturbance than the non-delinquent controls.

Contrary to theories of 'cultural' delinquency, the probationers in high-delinquency areas were just as maladjusted as those in 'good' neighbourhoods. But the non-delinquents were consistently less well adjusted the greater the density of delinquency.

It is suggested that delinquency is mainly an aspect of more general behaviour disturbance in the under-privileged quarters of large cities, and that this in turn is only one of several consequences of unfavourable ecology.

I. THE SOCIOLOGICAL-PSYCHOLOGICAL DICHOTOMY

There is a tendency for the disciplines of sociology and psychology to offer self-contained explanations of delinquency which are treated as antagonistic to one another. Cohen (1956), for example, hardly does justice to psychological studies in declaring that 'These theories have in common the idea that delinquency is a result of some attribute of the personality of the child, an attribute which the non-delinquent child does not possess or does not possess in the same degree.' In effect it would be hard to find a contemporary psychological theory which does not envisage all behaviour, delinquent or otherwise, not just as an 'attribute' but as the culmination of a history of interaction between the agent and his environment.

A similar dichotomy is apparent in the opening statement of the 'Radby' study by Jehcott & Carter (1954): 'There are, then', they say, 'two main points to be emphasized. The first is that there may be delinquency arising out of environmental or cultural circumstances in addition to delinquency arising out of psychological or pathological disorder.' Later they write, 'What we hope to show is that the bulk of the delinquents come from certain localities and types of households which present traits which are conducive to delinquency and traits which are social rather than psychological.' This statement seems to the writer to epitomize the confusion of thought responsible for the dichotomy. No complex of behavioural phenomena, such as can be observed in a family or a street, can be rightly described as either social or psychological; it is only their treatment—the selection of the observations, their classification and the concepts in the minds of the investigators which can be described as the one or the other. The sociologist deals with regularities of group behaviour, the psychologist follows the same behaviour as seen in the motivation of individuals. Any attempt to explain a behavioural phenomenon entirely in terms either of group reactions or of personal characteristics leaves large questions unanswered. Thus, Cohen, in advancing a sociological theory, has to elaborate it in the psychological concepts of guilt-feelings, shame, and frustration. Jehcott and Carter show how the social traits adduced by them have certain psychological effects

upon the children, notably anxiety and insecurity. For the complete understanding of how any social ecology can produce certain types of behaviour it is necessary to carry the inquiry into the psychological sphere in order to ensure that the responses adduced are consistent with more general observations about the mental and emotional processes. Moreover, only by psychological investigation can we explain variations of behaviour within culturally homogeneous groups, as when only a proportion of the juvenile population of a 'delinquent sub-culture' become delinquents. On the other hand we have to enter the sphere of sociology in order to understand those modes of behaviour which are culturally stimulated, and in order to follow the general regularities of interaction between the individual and his cultural setting. For example, an abstract conscience, or 'super-ego', explains little. It is a *psychological* fact that the individual has an urge to adapt his behaviour in order to become acceptable to the group; but the expectations of the group towards him are *sociological* facts. In short, the two approaches can never offer alternative explanations of delinquency or of any other behaviour; each discipline examines the dynamic of attendant conditions in its different aspects and phases, and so is complementary to the other.

II. THE 'DELINQUENT SUBCULTURE'

When they oppose sociological to psychological explanations it is natural that the proponents of each discipline should tend to exaggerate the extent to which the crucial influences can be described in their own terms and concepts. It is rather in the evaluation of the relative importance of pure cultural stimulation and of ecological influences operating through emotional and behaviour-disturbance that the present writer is inclined to differ from Mays (1954). His statement that, 'environment supplies both the stimulus and the content for delinquency and it is in the interaction of environment with individual psychological factors that the delinquency of any particular boy is to be understood', is unexceptionable. But it is more doubtful if, 'in the majority of cases delinquency can be regarded as a phase of normal development within a particular environmental setting', or whether 'in the present state of things a temporary period of delinquency is an unavoidable consequence of growing up in modern urbanized conditions...'. The crux of the question is whether, as he argues, 'delinquency is not so much a symptom of maladjustment as of adjustment to a subculture in conflict with the culture of the city as a whole'. In his study of the docks area of Liverpool, Mays was probably studying the marginal case in which such 'cultural' delinquency was an important factor making his statement as nearly true as anywhere. An Irish Catholic population of unskilled manual labourers could be expected to feel little sense of community towards a more well-to-do Protestant majority. The property of the out-group would not be sacrosanct, especially where goods in their passage through the dock area would be so easily stealable. But even in these conditions conducive to the formation of a 'delinquent subculture' only certain types of larceny were condoned. As Mays says, 'some crimes like breaking and entering are regarded as being in a different category. While not proscribed by the group they are regarded as dangerous and excessive... and little gangs who are known to break into buildings are usually avoided while engaged on such pursuits.' It is a commonplace that within every cultural group there is a condonement of certain classes of

behaviour which are legally prohibited, and the more remote the victim is from the group the greater the tolerance. Mays's study embraced a community in which the permitted illegalities were other than those ruling in more typical working-class communities. A large part of the difference of emphasis, between him and the present writer, in effect rests upon his rather special distinction between 'delinquents', as the individuals who 'succumb temporarily to the social infection of the delinquent area'—that is, who exploit the condoned illegalities of their culture—and the 'criminals' who commit the uncondoned offences. Mays, nevertheless, admits that even the 'cultural' delinquency is rendered more likely by adverse family environment. 'To sum up the argument . . .', he says, 'in certain urban areas with a tradition for antisocial behaviour, children tend to react to personal difficulties or to strained relationships in the home by delinquent behaviour. But where relationships within the home are not exceptionally bad or of long duration the amount of delinquent activity necessary for adjustment is limited and temporary.' One would have supposed that a certain amount of general behaviour disturbance would have been a concomitant of home difficulties severe enough to induce delinquency, but Mays claims that since these temporary delinquents respond to cautioning, fining and probation this 'is a proof that they are not maladjusted'. He is at pains to emphasize the psychological normality of what he calls the 'purely environmental delinquent' (meaning apparently in this sense cultural), and quotes a typical example in the boy, Arthur, who has taken part in a number of nefarious activities, albeit 'well-adjusted to his milieu'.

The empirical confirmation of Mays's thesis depends, consequently, upon objective data as to the non-delinquent behaviour of a group of city delinquents. If their delinquency is predominantly a matter of 'social infection' and of 'adjustment to a subculture' one would expect them to behave as normally adjusted persons when they were not indulging in the condoned illegal acts. That is to say, they would behave in a way that would on the whole further their own best interests while maintaining good personal relationships with others of their group. In particular, they would get on well, or at least, not have a hostile attitude to, their age-mates within their own cultural setting. The second condition for the validity of Mays's thesis—which admits family troubles as a facilitating factor—is that in the most delinquent areas, where the social infection is strong, one would expect to find that lapse into delinquency would occur with a lesser degree of family trouble, and hence few signs of general behaviour disturbance. Conversely, in the non-delinquent area, delinquency would only be resorted to under very adverse family conditions, with a high degree of concomitant maladjustment.

The issue hinges, then, on the type and amount of general behaviour-disturbance found among delinquents. Perhaps it should at this stage, however, be repeated that experimental verification would not consist in allocating the proportions of delinquents who become so for cultural reasons on the one hand and psychological reasons on the other. That delinquency is a response to unfavourable ecology cannot be doubted. What needs to be determined is to what extent delinquency is induced by direct cultural suggestion—as the theory of the delinquent subculture implies—and to what extent the adverse factors in the ecology produce a general behaviour disturbance, of which delinquency is only a part—one might say the small part of the iceberg which is visible.

III. THE GLASGOW EXPERIMENT

The total of Glasgow boys under 15 years of age (and so still at school) who were put on probation for the first time during 1957 were chosen as the subjects of the investigation. They represented about one-third of all boys of their ages who had been found guilty of offences against property or the person. To this extent they were a selected group; but the procedure for dealing with young people charged in Glasgow hardly allows time for the collection of the information needed to decide which cases should be awarded probation, the figure of one-third being an administrative 'ceiling' determined by the size of the probation staff. An analysis of the probation officers' reasons for recommending probation is given below.

To provide controls the head-teachers of the schools attended by the probationers were asked to take the nondelinquent boy anywhere in the school who was born on the same day or nearest thereto; if two more boys were equally near the birth-date, the one whose name was highest in the alphabet was to be chosen. The control group was thus randomly selected, except that they came from schools in which there was more than the average number of delinquents. This, as will be seen from the figures quoted below, tended to produce a bias towards poor social adjustment.

In respect of each probationer and control the class-teacher was asked to complete the Bristol Social Adjustment Guide (Stott & Sykes, 1956). This was done in respect of 415 probationers, representing 82% of the total group, and of 404 controls. In some cases the class-teachers would have known of the fact of probation. However, this could have had little influence upon the records which they completed of the boys' behaviour, as these left no scope for interpretation or rating by observers. The latter had only to mark, from a large number of items provided, those which had been noticed in the subject, and these items detailed concrete, objectively observable and reportable behaviour or instinctive expressions of attitude whose meaning is equally instinctively comprehended by all social animals within their own species. The validity of each of these items, in its setting within the schedule, had been previously tested (Stott, 1958). 161 had been found to be typical, at a probability of 100 to 1, of maladjusted and 49 of well-adjusted behaviour. The 'maladjusted' items were also grouped into syndromes, so that the type of behaviour-disturbance, if any, could be ascertained. Since each type of disturbance of given severity results in approximately the same global score, the mean of the latter could be taken as an index of the degree of maladjustment within a group. Because a child who is maladjusted is so in a certain way, and the 161 adverse items cover all types, the most severely maladjusted child could hardly score more than 60, and the highest recorded in the present survey was 51. For the present inquiry, to obviate circularity, the items concerned with stealing or damage to property in or out of school—namely: HA 11, HA 17 and XC 13—were ignored. Thus the comparisons reported below are in respect of non-delinquent behaviour only. The distribution of the scores for the probationers and the controls are given in Fig. 1. The limit of up to four adverse markings for the stable is to make allowance for errors and misunderstandings and for isolated, apparently abnormal behaviour which may in fact be normally motivated. The relative social adjustment of the two groups can be shortly summarized as in Table 1. Here again a certain amount of, possibly temporary, unsettledness is allowed for in

the 'normal boy'. The figure of twenty was taken as an arbitrary 'plimsoll line' for maladjustment. The percentage of 7.7 of maladjusted given thereby for the controls is within the range of the estimates and survey findings undertaken on behalf of the *Report Underwood Committee on Maladjusted Children* (1955).

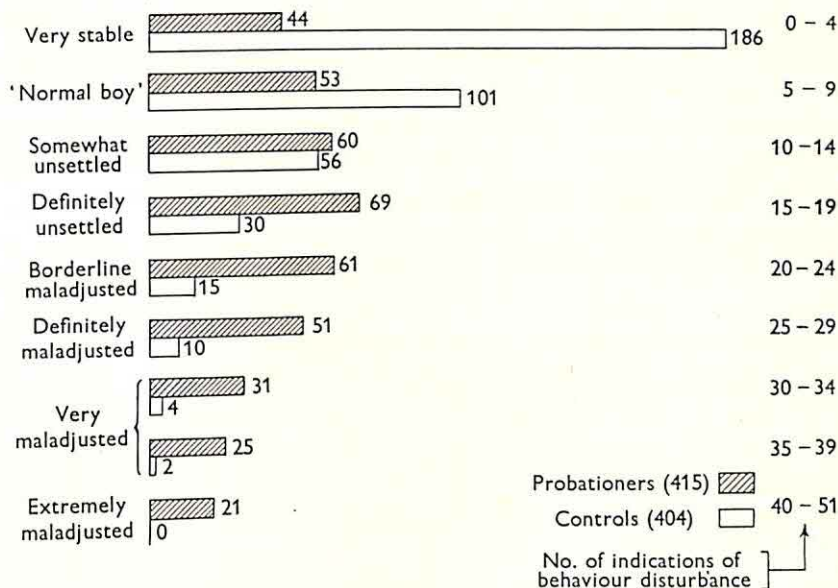


Fig. 1

Table 1. *Social adjustment of probationers and of non-delinquent controls*

	Criterion	Probationers (415)	Controls (404)
'Normal boy'	0-9 adverse markings	23.3 %	71.0 %
Maladjusted	20 or more adverse markings	45.6 %	7.7 %

That nearly half the probationers passed the 'maladjusted' score, and less than one-quarter fell into the 'normal boy' class speaks strongly in favour of the hypothesis that adverse ecology operates globally through general behaviour-disturbance rather than specifically through cultural suggestion.

Since the probationers were selected from among a much larger total of similar offenders it must be asked whether the prevalence of maladjustment among them might be due to a tendency of probation officers to recommend the more maladjusted or unstable offenders for probation. An analysis of the reasons given for their recommendations is reproduced in Table 2.

It is seen that instability of character only comes fourth in order of frequency, with some 10 % of the total reasons given. A detrimental family situation (reason no. 3) would also indirectly be likely to select more maladjusted children, but this category was a broad one including the ill-health of a parent, and the child living with other than its parents; moreover it only covered 16.4 % of the reasons. No more than eleven of the probationers had previously been referred to a child-guidance

clinic, and only one was at the instance of the probation officer. Thus it cannot be said that the probationers were selected to any significant extent because of their general maladjustment.

There still remained the possibility that the 23 % of comparatively well-adjusted probationers represented ordinary boys from the delinquent subcultures. If direct cultural suggestion were a significant factor we would expect to find that the probationers coming from these areas would include more of these well-adjusted boys than the probationers residing in good areas. Glasgow seemed a suitable place to test this hypothesis, for certain quarters of the city are traditionally known for their lawless tendencies; in fact it might be said that if delinquent subcultures were to be found anywhere they would be in Glasgow. The problem was how to demarcate these

Table 2. *Probation Officers' reasons for recommending probation*

	No.	%
1. Parents considered lax, negligent, unwilling to discipline	143	18.1
2. Boy unable or unwilling to keep away from undesirable companions	132	16.7
3. Family situation known to be detrimental	130	16.4
4. Boy considered of undesirable character	78	9.8
5. Father not at home and a man's guidance needed	66	8.3
6. Other children in family previously delinquent	65	8.2
7. Boy's repeated offences	48	6.1
8. Parents have set bad example, by delinquency, etc.	39	4.9
9. Gravity of offence	28	3.5
10. Other reasons	28	3.5
11. Probation awarded although not recommended	35	4.4
	792	

areas with fair accuracy and objectivity. This was complicated by two factors. The first was that the Catholic section of the city are fairly well distributed throughout the working-class quarters, and it was to be anticipated that the cultural differences between Protestant and Catholic would be more than purely religious. The second was that owing to extensive rehousing, certain of the old areas are changing their character, so that traditional neighbourhood boundaries could not be relied on. The electoral wards of the city were of course too large, and did not in any case correspond with cultural differences. By way of exploring the position the places of residence of all the probationers who attended secondary schools (12-15 years) were plotted by map-pins on a street map of Glasgow. The colours of these pins showed their schools. It was seen that the delinquent concentrations were marked, but were in comparatively small areas. These also seemed to correspond to the catchment-areas of certain schools. These latter therefore suggested themselves as the territorial units of the cultural patch-work. They had the additional advantage of taking account of religious, and therefore ethnic, differences, as the schools are denominational. For the secondary schools there remained an element of inaccuracy in that some of them, mostly the 'senior-secondary' (selective grammar schools) drew their pupils from a wide area; and a few pupils in others, for rehousing or other reasons, did not live in the localities of their schools. Consequently, where three or more of the probationers of the sample attending one school lived outside the normal catchment-area all probationers attending the school were left out of consideration on the assumption that the school would be culturally too heterogeneous. Where one or two probationers

were seen by the above-mentioned map to be living outside the neighbourhood of their schools, they were removed individually. The numbers remaining are given in Table 4 below. By these means the school catchments probably represented neighbourhood units which were culturally fairly homogeneous.

The next task was to classify these school catchment-areas by the degree of delinquency endemic within them. Since the number of boys on probation for the 1 year in question was not high enough to give reliable incidence for these small areas, the numbers put on probation for the previous 2 years in each school were added. Thus every school was given a delinquency index which consisted in the mean percentage of boy-pupils placed on probation for the first time during the previous 3 years. Apart from those which had none at all and were left out of account as providing no data, the schools were divided into four approximately equal groups according to their delinquency indices. In order to make these groups of about the same size different limits had to be chosen for primary and secondary schools. These are given in Table 3.

Table 3. *Classification of schools by mean percentage of probationers over previous 3 years*

Incidence	Primary (8-12 years)	Secondary (12-15 years)
Low	1.3 or less	5.4 or less
Next low	1.4-2.2	5.5-6.6
Next high	2.3-4.4	6.7-9.2
High	4.5 or more	9.3 or more

Table 4. *Mean maladjustment scores of probationers by cultural units classified by incidence of delinquency*

Delinquency classification	No. of probationers	Mean malad- justment score
Low	55	18.727
Next low	65	19.753
Next high	70	17.100
High	102	18.862
	292	18.613

For each of these school-based cultural units the mean maladjustment scores (number of adverse markings) were calculated. These are given in Table 4.

None of the differences in the mean maladjustment scores were significant. The greatest, that of 2.653 between the second and the third, compared with a standard error of the means of 1.978, gives a *t* of only 1.34; so that the somewhat higher mean of the next lowest and the somewhat lower mean of the next highest must be attributed to chance or to some unknown factor irrelevant to the hypothesis tested. The difference of 1.627 points between the first and third is, with a *t* of 0.77, still less significant. Any possibility that these non-significant differences represent very small real tendencies for probationers in the high-delinquency groupings to include more 'normal' boys must be ruled out by the reversal of the tendency for the highest-delinquency area. Here the mean level of maladjustment was surprisingly similar to that of the lightest area and to the total mean score. The evidence, in short, does

not thus far support the hypothesis of a significant cultural infection of the psychologically normal boy-population within the so-called delinquent subcultures.

There remained the possibility that the above classification of delinquents on the basis of the catchment areas of their schools did not reflect true subcultural divisions. It might, for example, be held that there were factors within certain schools which were conducive to delinquency. But since there were twenty-six schools in each of three groupings and twenty-five in the fourth this is extremely unlikely. Nevertheless, an alternative means of isolating the probationers within the 'delinquent-subcultures', which would be completely independent of their schools, seemed advisable. This was possible owing to the plotting of the places of residence of the secondary-school probationers, it being in addition feasible that this group, aged 12-15 years, would be more susceptible to the direct cultural infection of an area than younger boys. It was a question, therefore, of selecting all those who were members of the densest concentrations. Where five or more probationers of secondary-school age were found to live in a circle of a quarter of a mile in diameter this latter was designated a 'black spot', and the forty-seven probationers in question were made into a special group representing the most delinquent of 'delinquent subcultures.' Their mean maladjustment score was 18.297—very near those for the least-delinquency area and for the mean of the whole. Among this group one would surely have expected to find a significant tendency to normality among delinquents if it be true that within certain environmental settings delinquency is a phase of normal development.

To sum up the findings reported up to this stage, it has been seen that the probationers as a whole include about six times as many maladjusted boys as the controls, but they are equally maladjusted in whatever locality or cultural grouping they live. Since only a proportion of the maladjusted boys would be expected to become delinquent within any given locality, it is reasonable to infer that there must be a larger volume of non-delinquent maladjusts, as well as delinquents, within the high-delinquency areas. This would also be consistent with the hypothesis of an unfavourable ecology responsible for behaviour-disturbance, of which delinquency is the presenting symptom.

To test this latter possibility the above groupings of schools were also used for the controls, and the mean scores for maladjustment similarly calculated for them. As seen in Table 5, this had the effect of relating the level of non-delinquent maladjustment with density of delinquency within the four cultural groupings.

Unlike the parallel figures for the maladjustment of the probationers, there is here a clearly discernible trend: the incidence of behaviour-disturbance among non-delinquents rises consistently with the amount of delinquency within the cultural locality grouping. The increases in mean scores are not large enough to be significant between each step, but the difference between the lowest and the highest is so at a 100 to 1 risk of chance ($t = 2.631$), and that between the two lowest and two highest combined is just so at a 20 to 1 risk ($t = 1.971$).

The same trends show themselves equally clearly by a different treatment of the data. It was seen that a score of twenty or more was taken as the criterion for maladjustment, and nine or less for that of stability or quasi-stability. The intervening group, who scored between ten and nineteen, may be termed the unsettled.

Table 6 gives the percentages of probationers and controls who fall into each class, by cultural groupings. These are also shown graphically in the histograms of Figs. 2 and 3.

Table 5. *Mean maladjustment scores of non-delinquent controls classified by density of delinquency*

Delinquency classification of area	No. of controls	Mean maladjustment score
Low	55	5.218
Next low	63	6.904
Next high	63	7.365
High	108	8.148

Table 6. *Distribution of probationers and non-delinquents within social-adjustment classes (percentages)*

	Delinquency class of area				
	Low	Next low	Next high	High	'Black spot'
Probationers					
Maladjusted	45.5	49.2	42.9	40.2	36.2
Unsettled	27.3	24.6	28.6	35.3	42.6
Normal	27.3	26.2	28.6	24.5	21.3
Controls					
Maladjusted	3.6	7.9	9.5	9.3	—
Unsettled	12.7	12.7	19.0	27.8	—
Normal	83.6	79.4	71.4	63.0	—

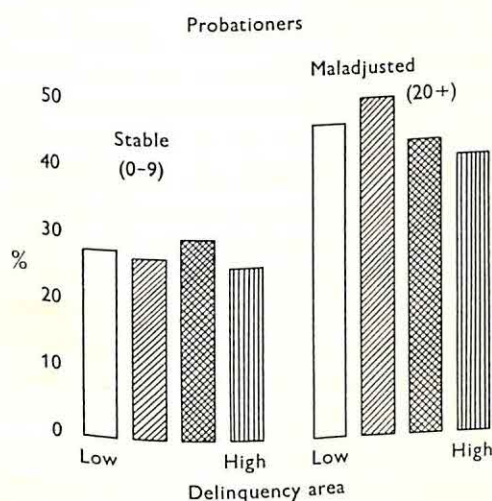


Fig. 2

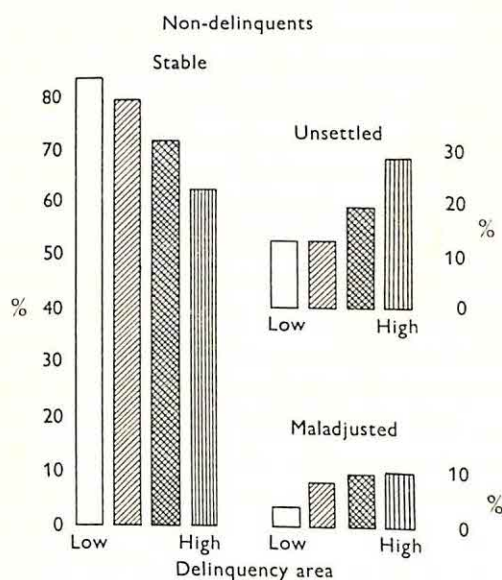


Fig. 3

It will be appreciated that the above method of classification, based upon the numbers meeting arbitrary criteria, allows more chance fluctuation than that of the calculation of mean differences. This probably accounts for the somewhat higher

percentage among the probationers of maladjusts in the next-low group, for which no meaning can be suggested. The small numbers of the black-spot probationers in each social-adjustment class probably also accounts for their slightly smaller proportions of both maladjusts and normals and rather more unsettled. It has been seen that these variations were not reflected in the mean differences quoted above. In any case, for the probationers no over-all trend as between the cultural groupings is discernible, except that those within the 'black spots' tend to be rather less maladjusted but more unsettled.

For the controls, on the other hand, the deterioration in social adjustment as between the lower and higher delinquency concentrations shows itself with equal consistency, and even more clearly. The incidence of maladjustment was over $2\frac{1}{2}$ times greater in the high- compared with the low-delinquency area, that of 'unsettledness' was over twice as high, while that for good adjustment within the normal range fell correspondingly.

Before any conclusions can be drawn from the above findings an explanation alternative to that of the hypothesis of ecologically induced behaviour-disturbance must be considered. Could it be that the social adjustment of the pupils, as reported by their teachers, reflects cross-cultural conflicts or differences in norms of behaviour, in other words that the stable child of the unacademic culture reacts unfavourably, or cannot adapt himself, to the norms of the culture to which the teacher belongs? Such tendencies do make themselves apparent to a very limited degree, for example, in the impatience of the low-culture child to be out earning money. But the Bristol Social Adjustment Guides, upon which the assessments were based, were validated upon samples of working-class children. In the course of this work it was indeed noticed that the relations of girls, but not boys, in secondary modern schools with their teachers tended to deteriorate somewhat during their last year at school. But the norms were framed so as to include within the normal all such minor deviations from optimum adjustment. A behavioural manifestation was only rated as unfavourable if its distribution as between the maladjusted and stable groups gave a risk of chance of less than 1 in 100. The type of maladjusted behaviour most characteristic of the delinquents of the present sample also tells against the concept of culturally induced 'maladjustment'. A prediction instrument was constructed from these data by the selection of those items which occurred at a minimum four times more frequently among the probationers than among the controls (Stott, 1960*a*, *b*). The syndrome which had the highest proportion of items (eight out of ten) meeting this criterion was that of 'hostility to other children'. If the maladjustment detected by the technique used was in fact a cultural phenomenon—the reaction against a foreign cultural intrusion—one would not have expected it to be expressed most strongly in anti-social attitudes towards other children of the same cultural grouping. In effect it represented the exacerbation of one or other of the instinctive executive reactions, such as would have been unwelcome in the ordinary citizen of any culture (Stott, 1956).

The hypothesis of culturally induced maladjustment, moreover, does not tally either with common experience or with evidence from particular communities. As regards the first, it is surely a matter of common observation that stable people react to non-critically unfavourable situations not by demonstrative hostility, but in a

way that will minimize their involvement; they avoid trouble. Faced with a teacher whose personality or ways they dislike, stable children become correct and avoid familiarity. Such reactions would not have scored adverse markings on the Bristol Guides. Nor can it be said that low-culture children are in general antagonistic or unadaptable towards the representatives of 'educated' cultures with whom they come into contact. It would make a fruitful study to observe systematically the adjustment of gypsy and bargee children to school-teachers. The experience of teachers in the Gloucester area, as reported to the present writer, was that these types and the children of fair-ground people were outstanding for their adaptability and friendliness during their temporary sojourns in one school or another—even though they often had to be hunted up by the school welfare officer to get them to school.

As regards any evidence from anthropological and similar studies of actual communities in Britain, one notes from that of Frankenberg (1957) in particular how in an industrial village community the cultural 'outsiders' are put into leading committee positions because their rulings can be accepted *without* the antagonism that they would evoke if coming from members of the in-group. Spinley (1953), describing the personality characteristics and social attitudes of a typical boy of a London slum, sums up one of the traits as, 'He says that a good leader should be authoritarian', adding, 'The children seem worried when allowed a wide latitude in their classes and clubs.'

Charles Russell, the London club pioneer who worked in an East End culture which was at that time—the early twentieth century—divorced from the mores of the dominant culture to an extent that it is difficult to grasp nowadays, reported no such cross-cultural antagonism to the intrusions of himself and his helpers. True to the practice of the English public-school culture, he even caned his club members for their misdemeanours, and they still came to his club! What he observed as the most characteristic of these 'rough lads'—defined by him as those who possessed no Sunday suits—was that they were 'the most human and pleasing of the types which make up the club' (Russell & Rigby, 1908).

Theories of cross-cultural antagonism as a source of delinquent or maladjusted behaviour are themselves likely to be projections of the individualist, aspirational *Weltanschauung* on the part of theoreticians who are members of the dominant culture. Within the traditional culture of industrial and rural wage earners the chief 'value' would seem to be the maintenance of good relationships with other members of the in-group, for economic and social security is based on interdependence. Aspiration, either economic or social, is shunned owing to the risk of ostracism. In Whyte's (1955) *Street Corner Society* the lack of any real desire to move out of the small group into a wider political field, despite the opportunity to do so, is difficult for the dominant-culture reader to understand. Havinghurst & Taba (1949) in their study of adolescents in a small American town, also report the conflicts which young people subscribing to the traditional, non-aspiring culture, face when they realize that advancement to better social positions may estrange them from their families and friends. Against these objective reportings from actual communities, in which cultural conflicts are seen to centre more around the prospect of moving out than around the points of cross-cultural contact, purely theoretical essays such as that of Cohen

carry little weight. It is noteworthy that the latter author quotes no evidence in support of his thesis that delinquency springs from the social frustrations felt by lower-class youths because of their inability to compete within the aspirational value system of the dominant culture. In summary, while the theory of 'cultural maladjustment' has a certain seductive appeal from the point of view of a rounded-off culture theory, factual corroboration for it has yet to be produced, this being all the more necessary in that it does not tally with the observations of those who have first-hand knowledge of the cultural situations in question.

IV. DISCUSSION

It does not seem feasible to explain the strikingly greater incidence of maladjustment among the probationers compared with their non-delinquent school-fellows as an outcome of cross-cultural contact. The same applies to the greater degree of maladjustment and unsettledness among the non-delinquents who lived in high-delinquency neighbourhoods. Nevertheless, these psychological findings have yet to be reconciled with the undoubted neighbourhood-distribution of delinquency. If, as suggested above, certain neighbourhoods offer an unfavourable ecology to a proportion of their youthful inhabitants, which results in a higher incidence of physical, mental and emotional maldevelopment, then the greater frequency of behaviour disturbance, including delinquency, could be seen as an aspect of this phenomenon. The evidence of physical impairment, the prevalence of mental defect in 'black spots', the association between mental defect and behaviour-disturbance point to environmental influences at the prenatal stages of development. For example, Anderson, Baird & Thomson (1958) have shown that the incidence of anencephaly (a malformation consisting of lack of a cerebral cortex) rises consistently with the deterioration of social conditions (per 1000 births in Scotland it is 1.29 for the crofting counties, 2.95 for the industrial central counties; for Aberdeen 1.7, Edinburgh 2.5, Dundee 2.9, Glasgow 3.1). The figures would correlate highly with the incidence of delinquency in the same areas. As regards unfavourable post-natal ecology, the studies of Spinley and of Jehcott and Carter quote abundant evidence to show that the manner of life within the low-culture areas is such as to deprive many children of family security, and a stable conflict-free atmosphere. In Spinley's words, 'The child of this city slum is exposed to so many textbook causes of maldevelopment that the problem is not so much why he displays the signs of disturbance which he does, but rather how he manages to continue as a functioning personality, as he also does.'

There is, in short, some independent evidence for the hypothesis, which is also consistent with the data quoted in the present paper, that the city-ecology proves unfavourable for human development as far as the least privileged sections of the urban population are concerned. The resulting maldevelopment takes the form of the physical, intellectual and emotional (motivational) impairment of a certain number of individuals. Of the last sort, delinquency, because of its nuisance value and the legal proceedings involved, attracts public and academic attention. But there would seem to be other aspects of behaviour-disturbance which pass comparatively unnoticed. And it would be expected, if the above hypothesis is well founded, that those who present one symptom of behaviour disturbance should be prone to others.

Hence the close relationship between delinquency and the psychologically classifiable types of maladjustment.

Just why and how the urban ecology proves unfavourable are complicated problems touching many fields of knowledge. There can be little doubt that adverse material living conditions—overcrowding, lack of sunlight, atmosphere pollution, overwork of the womenfolk, lack of money as far as large families are concerned—play a large part. But the widely differing proneness to breakdown within similar economic strata, as so well shown by the study of Jehcott and Carter, suggest that there is an important cultural factor. It would be an over-simplification to say that certain subcultures are simply more delinquency-prone, or delinquency-tolerant. Anthropologically, the concept of a culture as a complex of those attitudes, beliefs and approved behaviour which constitute a distinct way of life presupposes a degree of stability and biological success. This cannot be said of the London slum of Spinley, and hardly of the 'black streets' of 'Radby'. Nevertheless, in the latter we see the elements of a culture which is so widespread as to be almost the rule in primitive communities. Within the original ecology of the latter it is obviously successful, considering that it has provided the framework of village communities from the Neolithic phase onwards. But its success depends upon a firm network of mutual obligations and mutual dependence (Malinowski, 1950), and a framework of discipline both for young and old provided by approved social practice and taboo. In the city community, especially when, as with the Irish or Puerto Ricans, there is immigration direct from the village community, the supports provided by the system of mutual help are knocked away before the alternative pattern of individual independence can be substituted; the discipline of the traditional framework, upheld by the public opinion of a small group whose approval is the individual's paramount value, becomes forgotten and unenforceable in a loose-knit urban conglomeration. It is not, then, that certain subcultures are inherently delinquency-tolerant—which would be anthropologically unique and therefore unfeasible—but that, within an ecology for which they are ill-adapted, the cultural norms decay and lose their force, leaving nothing in their place. Since the biological *raison-d'être* of the phenomenon we term a 'culture' is the acquisition and handing on of a body of folk-wisdom, a recipe for successful existence, it follows that the absence of any culture, or the existence of one inappropriate to the contemporary ecology, threatens the loss of that high degree of individual security which the human cultures have achieved, in contrast to the biological hazards of animal survival. In detail, this is seen as an increase in material and hence in psychological stresses and strains, the overburdening of the psychic tolerance of individuals, and the greater frequency of personal breakdown and aberration.

REFERENCES

- ANDERSON, W. J. R., BAIRD, D. & THOMSON, A. M. (1958). Epidemiology of stillbirths and infant deaths due to congenital malformation. *Lancet*, i, 1304.
- COHEN, A. K. (1956). *Delinquent Boys*. London: Kegan Paul.
- FRANKENBERG, R. (1957). *Village on the Border*. London: Cohen and West.
- HAVINGHURST, R. J. & TABA, H. (1949). *Adolescent Character and Personality*. New York: Wiley.
- JEHCOTT, A. P., CARTER, M. P. & SPROTT, W. J. H. (1954). *The Social Background of Delinquency*. Nottingham.
- MALINOWSKY, B. (1950). *Crime and Custom in Savage Society*. London: Kegan Paul.

- MAYS, J. B. (1954). *Growing up in the City*. Liverpool: The University Press.
- RUSSELL, C. E. B. & RIGBY, L. H. (1908). *Working Lads' Clubs*. London: MacMillan.
- SPINLEY, B. M. (1953). *The Deprived and the Privileged*. London: Routledge.
- STOTT, D. H. & SYKES, E. G. (1956). *The Bristol Social Adjustment Guides*. University of London Press.
- STOTT, D. H. (1958). *The Social Adjustment of Children*. University of London Press.
- STOTT, D. H. (1960a). The Prediction of Delinquency from nondelinquent Behaviour. *Brit. J. Delinq.* **10**, 195.
- STOTT, D. H. (1960b). *A Delinquency Prediction Instrument*. University of London Press.
- Report Underwood Committee on Maladjusted Children (1955). H.M.S.O. London.
- WHYTE, W. F. (1955). *Street Corner Society*. Chicago.

(Manuscript received 17 August 1959)

A PROJECTOR FOR COLOUR DEMONSTRATIONS

BY J. R. TROTTER

*Department of Psychology, University of Edinburgh**

THEORY AND CONSTRUCTION

Light from the lamp filament passes through the condenser *A*, slides *B* (any standard 2×2 in. transparency) and *C* (a vertical strip marked with a linear scale 0-100 %), and forms an enlarged image of the filament at the aperture of the filter holders *D* and *E*. It is here modified by whatever standard 2×2 in. colour filters may be in use and then passes through the projection lens, which can be adjusted by the knob underneath to throw an image of the slides on to a screen. Holder *D* is fixed and any filter it contains colours the whole of the filament image. Holder *E* can take one filter or two, one above the other in a vertical plane, and it can be raised or lowered by turning the knob above. When it is fully raised the lower filter colours the filament image; when it is fully lowered the upper filter does this. At intermediate positions the upper and lower parts of the filament image are separately coloured by the two filters in any desired proportions.

The display on the screen has the geometrical pattern of the slides and is uniformly coloured by the filament image. The colours produced by the two filters in holder *E* are superimposed on the screen, giving an additive mixture of lights. Since the *E* combination acts on the light as already modified by the filter in holder *D* their joint effect is what is usually called subtractive mixture (an unfortunate but well established name).

While any transparency can be projected, a useful display is the C.I.E. colour diagram showing the spectrum locus on $x.y$ co-ordinates. On the basis of published data co-ordinates of the screen colour can be calculated for any combination of filters and any position of filter holder *E*, so that a curve may be drawn on the diagram showing the co-ordinates for any setting. There is an arm extending from the filter holder *E* carrying a pointer which moves along the scale on slide *C*; the scale and the pointer are displayed on the screen. The pointer indicates that $x\%$ of the filament image is coloured by the upper filter and, by implication $(100-x)\%$ by the lower filter. This indicates the appropriate point on the curve and specifies the colour of the display.

The lamp is a Mazda A 1/53, 750 W. The condenser, of about 3 in. focal length, gives a filament image about $1\frac{1}{4}$ in. square at 12 in. distance. This is a convenient size, allowing the 2 in. square filters to be firmly supported. A 3 in. diameter achromatic doublet of 12 in. focal length is used for projection. Compared with the more usual 6 or 4 in. focal-length lens used in projectors this lens allows the lamp to be placed closer to the condenser and gives one-half or one-third the linear magnification of the slides, providing a screen luminance at least four or nine times greater, for a given throw. The increased luminance is necessary since many filters have a transmittance

* Now at Department of Psychology, Yale University, U.S.A.

of less than 1%. A longer throw is possible without excessive magnification and dimming of the display so that the projector can be placed behind a large audience, rather than in the midst of it.

The mating edges of the *E* filters should be clean and flat, since dirt or varnish will cut off some light and a gap will allow unfiltered light to pass. In either case the screen illumination will be uneven. Theory demands that, for perfect performance, the filament image and the filters should coincide in a geometrical plane. Gelatin filters are usually thin enough to be treated as planes but filaments, and especially their magnified images, have considerable depth. As a result the screen colour is not always perfectly uniform. When the optical thicknesses of the two filters in holder *E* are not equal double images of the slides are produced. These technical deficiencies make the projector unsuitable for research purposes but, for demonstrations where the utmost precision is not required, they are outweighed by flexibility and simplicity of operation.

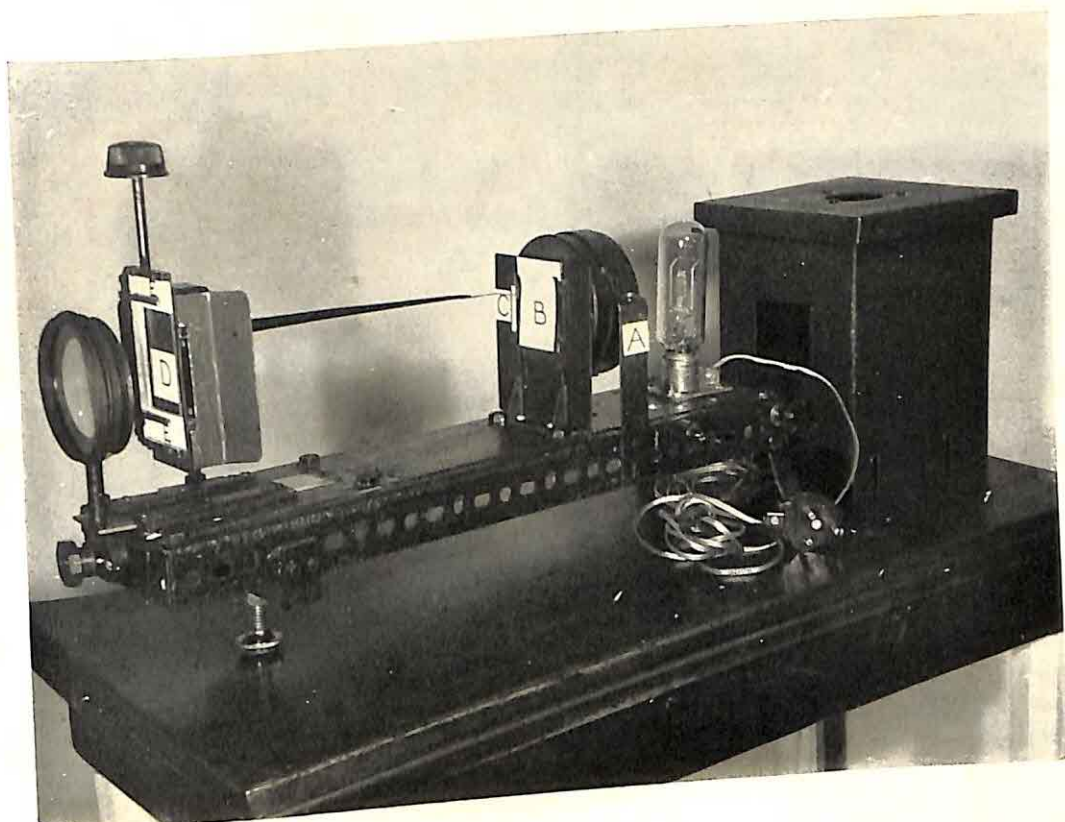
DEMONSTRATIONS

In an additive mixture the physical correlates of hue, saturation and brightness, viz. dominant wavelength, purity and luminance can be varied by leaving holder *D* empty and using suitable pairs of filters in holder *E*. For dominant wavelength use two differently coloured filters of similar transmittances. If the filters are very different saturation changes also, and if they are complementary white can be produced. For purity use any colour filter and a neutral filter of about the same transmittance.

To demonstrate subtractive mixing a filter is placed in holder *D* and another in holder *E*. Depending on the position of holder *E* the screen colour varies from that of the *D* filter to that of the subtractive mixture of the two colours. If the transmission spectra of the two filters have nothing in common the subtractive mixture will be black. If two differently coloured filters are in holder *E* the colour changes from the (*D* and *E* upper) subtractive mixture to the (*D* and *E* lower) subtractive mixture. The complexities of subtractive mixing can be illustrated by using in holder *E* two filters which look the same but have different transmission spectra. Though these give the same additive mixtures they give different subtractive mixtures.

All the phenomena of after-images which are commonly seen by a single observer using coloured papers can also be shown to a large audience, and may be very effective, thanks to the unusual brightness of the screen.

(Manuscript received 23 May 1959)



(Facing p. 172)

J. R. TROTTER

PUBLICATIONS RECENTLY RECEIVED

J. M. Charcot 1825-1893. His Life—His work. By GEORGES GUILLIAN. Edited and translated by Pearce Bailey. London: Pitman Medical Publishing Co. Pp. xvi + 202. 1959. 50s.

This is a straightforward, if somewhat anecdotal, account of the life and work of J. M. Charcot, together with a brief history of the Salpêtrière. Dr Guillian reminds us afresh of the brilliance and variety of Charcot's contributions to clinical neurology, as well as his possibly less fortunate incursions into the study of hysteria and hypnosis. He has also something to tell us about Charcot's personal characteristics, including the fact that he was devoted to animals and at one time kept a tame monkey which ate with the family. Charcot, like our own Henry Head, was a strong visualizer—a factor which may well have had something to do with his extraordinary powers of clinical observation as well as his interest in the visual arts. It is about time that someone looked afresh into the question of imagery differences and their possible relations to aptitude and interest.

This book provides a fascinating glimpse of the background of modern French neurology and of the man who did so much to create its characteristic pattern. Its translator, Dr Bailey, is an American neurologist who long ago trained with Pierre Marie, himself a pupil of Charcot, and has captured well the spirit of the French original. One may also commend the choice of Frontispiece: It is a sketch showing Charcot, in top hat, apron and galoshes (looking for all the world like the celebrated Dr Caligari, of filmic fame) gazing with rapt attention at a human brain held between his outstretched hands.

O. L. ZANGWILL

Free Associations: Memories of a Psycho-analyst. By ERNEST JONES. London: Hogarth. 1959. Pp. 264. 25s.

As his son describes in the brief Epilogue he contributes to *Free Associations*, Ernest Jones had started to write an autobiography in 1944, and had covered the period to 1918, the year before the foundation of the British Psycho-Analytical Society. It was at this point that he decided to turn his attention to a biography of Freud, a task for which his identification with the development of psychoanalysis from its earliest beginnings particularly fitted him. When, in 1957, Jones was able to return to his autobiography, he decided to revise what he had written rather than attempt to complete it.

Dr Glover's Obituary notice (this *Journal*, 49, 1958, 177-81) renders a further appraisal of Jones's role in the history of psychoanalysis unnecessary. What makes the *Free Associations* such fascinating reading is the opportunity it affords of seeing in the youth and young man the seeds of the determination and integrity which later made him so successful an exponent of an unpopular doctrine. Jones makes no claim to literary merit: in one of his many references to the influence on him of his friend Wilfrid Trotter (author of *Instincts of the Herd in Peace and War*), he writes: 'If I have learnt to avoid some of the grosser faults commonly exhibited by scientific writers it is entirely because I took that influence to heart.' Nevertheless, many of his own comments have an epigrammatic directness, which, coupled with their wisdom, might well qualify them for incorporation in any psychologist's *credo*.

B. SEMEONOFF

The Social Psychology of Groups. By JOHN W. THIBAUT and HAROLD H. KELLEY. New York: Wiley; London: Chapman and Hall. 1959. Pp. xiii + 313. 56s.

This book provides a new set of concepts for the analysis of small group behaviour, and interprets previous studies in these terms. The experimental work in this field has become diverse and fractionated, being inspired by a variety of miniature theories. There is thus a real need for concepts and hypotheses capable of co-ordinating this rapidly disintegrating field. The authors

of this book are well equipped to undertake the task: each has carried out several important small group experiments, and they collaborated over the review of the field in the Lindzey *Handbook*. They wrote this book while at the *Center for Advanced Study in the Behavioural Sciences* in 1956-57, where they were able to consult with Newcomb, Festinger and other small group experts.

The approach is novel in several ways. Two-thirds of the book is concerned with the analysis of two-person groups, and these ideas are then extended for larger groups. The two-person situation is analysed by means of a games theory model: each person has alternate behaviour sets in relation to the other, forming a payoff matrix. The payoffs are the subjective evaluations of outcomes in relation to 'comparison levels'—what the subject expects from the relationship. Probabilities of behaviour are said to follow the payoffs of alternate behaviour sets. No direct evidence for this basic model is given, but previous studies are shown to be consistent with the implications of it. For example, research in sociometry shows that people are chosen as friends who are most rewarding at least cost.

In this theoretical analysis, the authors use games theory much as Lewin used topology: the basic model is used as a way of representing social situations, by the abstraction of certain key features; on the other hand, no use is made of theorems from the theory of games; no rules are given for finding out in advance what the payoff matrix is. Consequently there is no concrete deduction of empirical consequences, and it is hard to see how the theory could be refuted. Extensive empirical material is cited, but the interpretations given are often highly inferential, as the experiments did not measure payoff matrices or comparison levels. Nevertheless, as with Lewin's theory, the analysis often leads to a new way of looking at old problems—the chapters on Tasks and on Interdependence in Larger Groups are cases in point. The analysis also leads to clarification of concepts and useful new distinctions—the attractiveness of a group and a member's dependence on it become separated, just as Lewin was able to distinguish different kinds of conflict situation.

The authors specifically reject the approach to small groups via motivation and learning, and as a result omit the growing literature on individual differences in social behaviour. This is perhaps a field where the integration of psychometric and experimental approaches, together with the application of existing knowledge about personality dynamics, learning and motivation, could well prove fruitful.

MICHAEL ARGYLE

Family Worlds. By ROBERT D. HESS and GERALD HANDEL. University of Chicago Press. London: Cambridge University Press. 1959. Pp. xiii + 306. 37s. 6d.

The blurb promises a new theory of family life, but the authors are more modest, claiming merely to provide a conceptual framework for the study of 'non-pathological family interaction and emotional structures'. In other words, their aim was to find a way of ordering the many bits and pieces of behaviour making up family life so as to bring out the coherent underlying patterns. One of the key concepts, for instance, is labelled 'congruence of images' and concerns the way in which each family member views and is viewed by all the others, the function of such images in the interaction process, and the kinds of changes these images undergo.

This promising start is followed by the 'psychological portraits' of five families, constituting the bulk of the work, in which the organisation of the material is guided by the concepts. These families were selected from a total of thirty-three, studied by means of interviews and several projective techniques. As indicated by the chapter headings (e.g. 'Flight from Insecurity', 'The Dynamics of Disconnectedness') each family illustrates a particular mode of mutual adaptation.

Whilst the portraits are sketched with sensitive skill, the present reviewer was assailed by an occasional doubt: the pattern sometimes appears too neat and rounded, and one senses an effort to press some of the responses into the proper mould. Here is an example: 'When Alice is asked how she feels about her father's job, she expresses her contentment with it in terms of its being a satisfactory spatial solution.' The answer which prompted this curious statement was as follows: 'I think he has a nice job. I think he likes it. He has a nice office and building.' Perhaps it may have been her interpreters, rather than Alice, who locomoted through the looking-glass. Apart from this reservation, the authors do offer a valuable new perspective, especially in

their concluding chapter where the various threads are drawn together. Thus they argue convincingly that discussions of the nature of authority in the family in terms of a democratic-autocratic dichotomy are apt to be misleading, since they tend to focus on limited and often superficial aspects of highly complex processes. What is perhaps most commendable is the valiant striving to hold an even balance between individual personalities and the family group as a whole. This is of course what social psychologists are always supposed to do, but in this sphere the gap between aspiration and performance is frequently wide.

G. JAHODA

Reflexes to Intelligence: A Reader in Clinical Psychology. By SAMUEL J. BECK and HERMAN B. MOLISH. Glencoe, Ill.: Free Press. 1959. Pp. xix + 669. \$8.50.

This is a big book, not only in physical size but in its aim and achievement. It presents seventy-three contributions selected from individual authors and is divided into seven sections. As the time-scale of the book is developmental these sections represent the seven ages of clinical psychology, with the important difference that, whereas man in his seventh age is decrepit, clinical psychology at this stage is presented as a vigorous discipline with an active future.

'Readers' of this sort seem to be current fashion in publications in clinical psychology. Up till now this reviewer has tended to regard them as being of the same ilk as Christmas annuals, bumper-books and the like, and valuable only if they chance to contain an article of current or classical interest. Editors of such collections are to be considered as mere compilers. This book is, however, different. Beck and Molish have been active as editors and, by integrating the selections, have produced a book which does more than supply a number of articles on clinical psychology and related disciplines, with the conviction that clinical psychology rests on a firm foundation of creative scholarship and, perhaps to the surprise of those who consider clinical psychology to be a collection of techniques, they prove their point.

In the first section they suggest classical, biblical and Shakespearian origins of clinical psychology. This is amusing and of novelty value but the real worth of the book blossoms at the beginning of Section II, 'Sources and Foundations', with an adequate outline of Darwin's contribution to psychology produced by blending three chapters of his 'Descent of Man'. The remaining sections are entitled 'The Quest for the Whole Person', 'The Measure of the Mind', 'Disciplines in Interaction', 'Humans in their Social Context' and the final prognostic section on 'Today's Theorists, Tomorrow's Realists'. It is interesting to note that the book starts with Darwin and ends with a biologist, von Bertalanffy, in whose thinking the editors see the development of the next stage of the description of the whole human personality.

In addition to the seventy-three separate articles which provide the broad coverage of differing viewpoints in each area, annotated auxiliary reading lists are provided, extending to about 2000 references. If this book had an index it would have to be two books!

This volume is certainly to be recommended to clinical psychologists. It is instructive for the psychologist beginning his postgraduate training and is informative and reassuring for the more experienced clinical psychologist who may in his darker hours have wondered about the scholarship of his discipline.

R. M. MOWBRAY

Concept Formation and Personality. By Z. P. DIENES. Leicester University Press. 1959. Pp. 82. 18s.

This brief monograph is an interesting and exciting, but not, in the reviewer's opinion, a very successful attempt to link concept formation and personality. What the writer has done effectively is to design a series of concept formation tests, and to administer these, as well as some personality tests, to groups of boys and girls. An attempt was then made to link up the two fields by means of canonic analysis. Dienes justifies the use of this technique in terms of the lack of any hypotheses in the field, and because he seems to consider that the results which can be obtained in this way bear some relation to Lewin's theories. Some relationships are found between cognitive and personality measures and these are discussed in some detail. The obtained structure for boys and girls appears to be different and this also is discussed.

The interest of the procedure and of the results is diminished by certain obvious objections. It is difficult in many cases to assess the statistical significance of the results as 'no ordinary significance tests are available for canonic correlations'. More important, however, are criticisms of the personality tests used. In the main these are projection tests, particularly Raven's controlled projection technique, and the author, although apparently informed of the uncertain validity of the results, appears to accept and interpret them at face value. The justification for this is very difficult to see. What does a score on 'anxiety' on such a test mean? What kind of criterion prediction does it make possible? What is its status in clinical and experimental psychology? As long as answers to these questions are not known, the demonstration of a relationship between test result and any other factor or construct is incapable of being discussed on a rational basis, or to form a fruitful subject of argument. Terms like 'anxiety' are almost entirely meaningless; as Cattell has shown, two experienced psychiatrists rating the same group of patients for 'anxiety' will show a correlation of only about 0.2. It is unlikely that such projective tests as were used by Dienes would show a correlation even as high as that.

While, therefore, this monograph is by no means free of faults, it does, nevertheless, show evidence of original thinking and makes use of methods which, with some development, could be very fruitful indeed. It should certainly be read by all those interested in concept formation and its relation to personality variables.

H. J. EYSENCK

Rorschach Responses of Elementary School Children: a Normative Study. By NETTIE H. LEDWITH, Ph.D. Pittsburgh University Press: 1959. Pp. xi + 185. \$4.00.

Adolescent Rorschach Responses. By L. B. AMES, R. W. METRAUX and R. N. WALKER. New York: Hoeber. 1959. Pp. xiii + 313. 63s.

All who use the Rorschach technique with children must be conscious of the lack of reliable norms for various ages. Both the above books represent serious attempts to supply this deficiency.

Dr Lewith reports the results of a longitudinal study of 138 children, to each of whom she administered the Rorschach test yearly from the ages of 6 to 11. In addition, two control groups aged 7 and 11, respectively, were tested once only, thus providing a check upon the factor of test-familiarity in the other children. The total number of Rorschachs given by the author in this normative study is nearly 1000, an achievement all the more impressive because of the obviously high level of test administration. Dr Ledwith wisely points out that although children tend to show Rorschach patterns that can be recognized as typical of different age-levels, there must be many deviations due to individuality without being in any way pathological. She gives a number of illustrative records, with brief but vivid case notes. This is essentially an informative book, in which fact and inference are never confused. The author seldom draws conclusions, but her data are clearly set out, and her comments are always lucidly to the point: they appear to be based on an excellent understanding of children, together with skill in the method employed.

Dr Louise Ames, with her colleagues at the Gesell Institute, cover the years from 10 to 16 under their title *Adolescent Rorschach Responses*, 100 children being tested at each of the seven age levels. The study is semi-longitudinal in character, single records being given by 271 subjects, and the remainder provided by 127 who were tested more than once. The data obtained are analysed in unusual detail, including interesting information as to the content of determinants. Inferences are freely drawn as to personality patterns considered as typical of each age, with colourful imagery and somewhat arbitrary conclusions. Two case studies are given for each year, one of a boy and one of a girl.

Since these two developmental studies overlap at the 10- and 11-year levels, it is possible to compare results for these ages. Although there is some difference in scoring, there are sufficient scorings in common to make comparison practicable in these cases. When this is done, there appears unhappily to be a wide divergence in findings. Children of 10 years, for example, give 27% of W in the Ledwith study; 51% in that of the Gesell team. Other comparative figures at the same age are F% 44 and 62; FC% 1.9 and 0.3. The use of M is more frequent in the Pittsburgh (Ledwith) sample, a surprising result, since the mean I.Q. of this group was considerably lower than that of the Gesell Institute, viz: 101.7 as against 115.9.

It may be that some of the discrepancies apparent are due to examiner-difference, an hypothesis supported by the contrasting average number of responses given. For the Ledwith samples of 10- and 11-year children these are 30 and 28, respectively: for the same ages in the Ames groups, 19.2 and 19.5. These differences cannot be accounted for by test-familiarity in the Pittsburgh children, for their 11-year controls, who were tested once only, have an average response number of 26. It has to be remembered also that of the Gesell children the majority were retested.

Both these publications provide valuable material, which should repay further study. Perhaps the main conclusions to be drawn from them in the meantime are first, that while developmental patterns are observable in the Rorschach, (as Ledwith emphasizes) human personality refuses to be pigeon-holed. Secondly, it appears that, in childhood especially, results are likely to vary according to the total two-body situation of the test.

THEODORA ALCOCK

Die Helldunkeldeutungen im Psychodiagnostischen Experiment von Rorschach.
(Shading interpretations in the psychodiagnostic experiment of Rorschach.)
By HANS BINDER. Bern and Stuttgart: Huber. 1959. Pp. 128. Swiss Fr/DM. 10.

The author, like Rorschach himself a Swiss psychiatrist, has been mainly responsible for developing the interpretation of chiaroscuro responses in the Rorschach test. His extensive work, first published in 1932 in the *Swiss Archives of Psychiatry*, is now made available as a monograph by Rorschach's original publishers.

In his discussion of the shading responses as an expression of dysphoria, Binder aptly calls attention to the place the later takes in the affective life of neurotics, and in particular that of psychopaths, in whom it assumes a central role. In so far as the technique can throw light on such finer points of psychopathology as this, sceptics of the value of the Rorschach are to some extent answered.

STEPHEN KRAUSS

Studien zur Testwissenschaft. By GERHARD CLOSTERMANN. Aschendorff, Münster, Westfalen. 1959. Pp. xx+363. DM. 28.

This book is devoted entirely to a discussion of results obtained through the use of a 'Draw-a-Man' Test. Work on this test had been started shortly after the end of the First World War, and had been going on steadily in the writer's Institute for a large number of years. There are twelve standard types of drawing which constitute a scale; the drawing made by the subject is rated as being at one of the twelve clearly defined points of this scale. The writer gives very exhaustive details about the performance of school children on this test, and the relationships of test results with school performance, social status, etc. He also discusses at great length the statistical problems involved.

There is little in this book that will surprise readers familiar with the Goodenough Test. The amount of detail given in the drawing correlates with school ability, with parental status and with other known functions of intelligence. The book is written in a manner which makes reading maximally difficult; indeed, it is not easy to conceive of anyone willing to plough through the lengthy discussion and the many tables for the very small return in new factual information that is likely to result. The statistical treatment, while reasonably orthodox, is curiously old-fashioned. As the writer claims to have a scale, one would have expected him to use some kind of scalogram analysis. Instead, he has restricted himself to the analysis of 2×2 tables and other similar simplified methods. In Germany, where quantitative psychology has but an insecure foothold, this may be a welcome change from the usual purely qualitative discussions; for more sophisticated readers, something more complex would have been desirable.

H. J. EYSENCK

An MMPI Codebook for Counselors. By L. E. DRAKE and E. R. OETTING. Minnesota University Press; London: Oxford University Press. 1960. Pp. 140. 30s.

The purpose of this book is to help a counsellor to 'form hypotheses about a relatively normal client on a basis of his MMPI profile pattern'. To this end a system of coding is offered, less complex than that set out in the *Atlas* of Hathaway and Meehl, but including ratings on the masculinity—femininity and social introversion—extroversion scales. Standardization of the Codebook was based on work with more than 4000 counselling cases at the University of Wisconsin, which on *prima facie* grounds suggests high probability of validity for a student population. Lacking experience, one cannot form an opinion of the actual validity of the hypotheses, but it may be worth recording that in an informal try-out with about a dozen psychology students two apparently contrasted female profiles each evoked the single hypothesis 'mother conflict'. By contrast, certain other profiles seem to be associated with such a wide variety of hypotheses that one feels one has arrived nowhere. One is left with the impression that this is not yet the handbook of MMPI diagnosis one had been hoping for. A need for something more universally applicable than this book and less cumbersome than the *Atlas* seems to be indicated.

B. SEMEONOFF

Information and Error. By SOLOMON DIAMOND. New York: Basic Books, Inc. 1959. Pp. xii + 307. \$5.00.

The author, who gives the subsidiary title 'An Introduction to Statistical Analysis', has succeeded in writing a pleasant book. The style is lucid and humorous and the illustrative examples—mostly of a psychological nature—are varied and interesting. But in common with most textbooks written in a light vein the book does not always succeed in combining informal language with exactness of expression. The title itself affords an example. It would appear (p. 6) that when the author uses the word 'information' he has *communication theory* in mind. But in communication theory the word 'information' refers to *uncertainty* and this is surely not the meaning of the term suggested by the title.

However, after this unfortunate start—and some confused notions about 'measurement' at the beginning of chapter 2—the book settles down to a discussion of introductory concepts in descriptive statistics and inference. Chapter 2 deals with measures of location. Chapter 3 is about sums of squares and at this stage the correlation ratio is (unwisely) introduced and reference to it should be omitted on first reading. Next come chapters on Variance—including a misleading section on 'partitioning variance'—Probability and Sampling Variance. Simple tests of significance follow, and they are tied up with variance ratio tests in chapter 8. The clarity of the latter chapter is ruined by an overdose of sigma-terminology.

In the next chapter the link between the large sample formula for comparing two proportions and the chi-square test as applied to fourfold tables is demonstrated. The 'continuity correction' is then introduced; but since the latter correction is always advisable it would now be well to forget about the earlier formulae.

As might be expected the author has good chapters on the correlation coefficient: its limitations too are suitably emphasized, and in this respect the diagram on p. 196 is effective. Next comes an unhelpful chapter on factor analysis which employs geometrical methods which are confusing and computing methods which are inadequate and outmoded. Finally, there is a chapter of bits and pieces entitled 'Handling Non-normal Data', in which the Poisson distribution fraternizes with a haphazard selection of ranking methods and alternative non-parametric tests.

It appears fair to conclude that while this book may appeal to the statistically lame student, who will find the pill lavishly gilded, the able student will be quick to turn elsewhere for a more precise—and more profound—introductory account of the topics discussed.

A. E. MAXWELL

Individual Behaviour and Group Achievement. By RALPH M. STOGDILL. New York and London: Oxford University Press. 1959. Pp. xi + 352. 40s.

Dr Stogdill has produced a closely argued theory and draws on a wide range of experimental evidence for its exemplification. The structure of the book is ingenious; he presents a very general indication of the theoretical elements, and then in each chapter he develops an aspect of his theory in the first half and gives the empirical evidence in the second. Finally he recapitulates, and the timid or harassed reader can confine himself to the final summing up, though much will be lost if this is done.

The outline of the theory is fairly simple. 'A group', says Stogdill, 'is regarded as an input-output system. The inputs are the performances, expectations and interactions of the group members'. These are the bare bones; in practice the input of some groups will include such items as raw materials, finance capital, information and so on. The basic unit of group formation is interaction, which clearly involves performance, though an individual may 'perform' alone from time to time in the interests of the group in which he is functioning. The third input factor provides the driving force and stability of the group. This is 'expectation', defined as 'readiness for reinforcement'. It is a 'function of drive, the estimated probability of occurrence of a possible outcome and the estimated desirability of the outcome.' The word, therefore, has a slightly different meaning from the one it has in ordinary discourse, and this unfortunately is not made explicit. On the one hand, a member behaves as he is expected to behave and as he expects others to respond; this is its consolidating aspect. On the other hand, an element of seeking is introduced. A man expects a certain outcome to be satisfying because it has proved so in the past, and now he is 'ready', in the sense of being 'all agog' for reinforcement (not merely accepting routine experience). In both senses his expectations will be influenced by probability; in the second sense an outcome may be so desirable that he will be 'all agog' to get it, however unlikely he is to achieve his aim.

We have, then, a net-work of interactions regulated by the 'expectation' of conformity, and by the 'expectation' of the participants that desired outcomes will result from their labours. Moreover, the system is 'open' in the sense that any individual participant can be replaced by a newcomer without the system losing its identity. The next item is 'structure'. For Dr Stogdill 'structure' involves differentiation of position within the system, such that members holding certain positions are not interchangeable with members holding certain other positions. 'The factor of reduced interchangeability', he says, 'may be regarded as the ultimate criterion of the difference between an organized and an unorganized group.' He seems to take this line because he wants to say that structure derives from personality differences between members; it is this *au fond* that makes them non-interchangeable. Thus when members are interchangeable we have a group of 'minimum structure'. Certainly most groups develop a structure in Stogdill's sense. There is, indeed, evidence that experimental groups tend to become structured, and are more efficient when this happens. Equally obvious is the fact that many groups are as it were, pre-structured in the interests of efficiency. Whether non-interchangeability is basic is debatable. You might have a tug-of-war team made up of men equally strong and equally skilful who could take turns at any position, and yet one would hesitate to say that it had 'minimum structure'.

The position up to date may be given in Stogdill's own words, by way of illustrating his style. 'The actions and reactions underlying the structures of positions established by interaction provide the behavioural instrumentalities for the confirmation of a system of mutually reinforced expectations which define for each position a set of predictive prescriptions relative to performances and interactions that accomplish group goals and maintain group structure'.

The next concept to be introduced is 'status'. The *status* of a position defines the degree of freedom granted its occupant in initiating and maintaining the goal direction and structure of the system in which the position is located. This is reminiscent of Homans. The leader is the person with greater freedom of action and greater scope for initiating interaction in others.

The group with its structure of positions with varying status is an abstraction. In every group there is an aim, and if the group is effective the members know the general function of each position and the general conduct required of its incumbent. A group, however, is always confronted with particular problems and the 'role' to be played by the incumbents will vary from group to group with the same structure, from time to time in the same group, and also in accordance with the characteristics of the incumbent himself. Thus the 'role' in this theory, with the

responsibilities and authority it carries, is linked with the predicament of the group and the personality of the player. It is thus possible to contrast the 'role structure' of two groups which have the same formal structure of status and function.

Such are the input variables and the intervening variables of structure. What of the output? This is analysed into three related features: productivity, integration and morale. The group produces a product—not always, of course a material one—the interacting and the satisfaction (or otherwise) of expectations leads to an accepted (or disliked) structure which can (or cannot) resist strains, while the lack of 'restraint' from within or from without gives rise to a state of morale, high or low. Stogdill carefully distinguishes between 'integration' which is associated with satisfaction, and 'morale' which is concerned with free-flowing enthusiasm. Now the interesting point is made that with the *same* input an increase in one of these aspects of 'achievement' is at the expense of the other two. In particular energy spent on increasing integration is liable to lead to decreased production. *Some* measure of integration is, of course, required, so is *some* degree of enthusiasm or interest, but the organizer must keep a balance between all three if the group is to carry out its productive purposes.

The details of the theory consist of variations on the main theme: how, for example, do the expectations of high status people differ from those of low status people? How do these two classes of people communicate? What difference is made if roles are tightly defined as compared with a loose definition? How do the different expectations of members affect their mutual expectations? And so on. This is where the experimental evidence comes in.

It will be seen that, in a broad sense, this theory might be taken as a further elaboration of the theoretical system propounded by Homans. It has the merit of providing a basis for further research, and a unifying framework within which to bring together a variety of very different empirical observations. To assess its value more fully one would have to spend a considerable time thinking in its terms, but at first glance one can certainly say that it looks promising.

The manifest function of Dr Stogdill's book is the exposition of his theory; it has the latent function of bringing together an enormous amount of experimental material from 794 articles and books which gives it a use-value of a no less welcome order.

W. J. H. SPROTT

White and Coloured. By MICHAEL BANTON. London: Jonathan Cape. 1959. Pp. 223. 21s.

This book is not easy to review because, although it contains many penetrating comments on white-coloured relations in Britain, it tends to leave the reader with the feeling that he has heard most of it before. The most valuable part seems to me to lie in highly perceptive comments which, as far as I know, have not been made before. The following quotation illustrates this. 'In Britain the white man is free to avoid the coloured man. . . . But the coloured man is not free to avoid the white man: he is dependent upon him for employment and social esteem.' The major criticism of this book is concerned with method because, in my opinion, the writer has attempted to answer psychological questions by sociological techniques only. The book begins with a set of queries of which the first will serve to illustrate this criticism. It goes as follows. 'Why should coloured people so often be shabbily treated when the vast majority of individual Britons are favourably disposed towards them?' The first criticism here is that what the vast majority of Britons think cannot be deduced from a sample of 300. The second is that though white subjects have said they are favourably disposed, they might either be consciously or unconsciously lying. Do, in fact, the people who say they are favourably disposed to coloured people, really act as if they were? This can only be ascertained by the behavioural test of observation and intensive field work.

An important part of the book is concerned with what Mr Banton calls the 'unspoken code' of the British, especially in relation to foreigners. Quite a bit of this 'code' could be the difference in manners and approach to strangers that exists between quite a lot of cultural and subcultural patterns. For instance, the clumsy attempt at getting to know coloured people by asking 'where do you come from?' may not be patronizing but an attempt to get on friendly terms by using the English gambit of placing or finding an identical interest with the person introduced. An analogous example is that in some European countries it is very bad manners to ring up before making a visit. Mr Banton suggests that the 'unspoken code' is stronger among higher than

among lower class groups. He suggests therefore that it is easier for a coloured man to be accepted in lower groups. It is difficult to say at this stage whether there really is enough evidence for this statement. Like so many other phenomena in social science, it cannot be assessed on its own. As most of the coloured people are working class they probably would not be taken as visitors to middle-class English homes. My own experience in Liverpool, too, has been that the lowest group, as exemplified by Ship Street, do not take non-relatives of any colour into their homes.

On the positive side, the distinction between prejudice and discrimination is a useful one. It does help to distinguish the behaviour which is mainly emotionally based from that which is mainly economically based. There is, however, a qualification I should make here. If, for example, a landlady turns down a prospective coloured lodger because she knows she would lose her other tenants, although she would be willing to take him herself, she may be discriminating, but her tenants would be prejudiced. Psychological and sociological frames of reference are frequently too interdependent to be so easily separated.

The least challengeable part of the book is the chapter rather regrettably called 'Ten Commandments of Inter Group Relations'. This consists of ten propositions about how to improve coloured-white relations. For example, 'clumsiness in interpersonal relations spoils the best intentions'. These are excellent, and should certainly prove invaluable aids to people who have to employ, work with, or be employed by, people of a different colour.

MADELINE KERR

La Genèse des Structures Logiques Élémentaires: Classifications et Sériations. By JEAN PIAGET and BÄRBEL INHELDER. Neuchatel: Delachaux et Niestlé. 1959. Pp. 295 28s.

Those who are at all familiar with the earlier work of the authors may be inclined to wonder at the title of this book on the grounds that precisely this topic has been discussed more than once already. The point of the new study is that, though the activities of classifying and serializing have indeed been discussed in a variety of connexions—such as the development of notions of time and number—and have been given a prominent theoretical role, they have not been studied directly and in their own right. The authors are setting out to fill what they see as an important gap in their work so far.

The book is divided into the expected three parts in the expected order: an introduction, a middle section reporting experimental work and discussing detailed findings, and a chapter drawing general conclusions. What is unexpected is that the authors, in a brief preface, advise the reader not to follow this order but to begin with the conclusions. He should then, they say, dip into the central chapters to see how the conclusions are justified, and finally he should turn to the introduction if, and only if, he feels constrained to read the work in its entirety. My first duty as a reviewer is to report that I tried to follow this advice and—rash as it is to disagree with the authors when they insist on a point such as this—was not finally convinced that it is good. Certainly, as the preface says, the main notions of the work are expressed in the brief space of the concluding chapter, and it may well seem to those who took part in the development of these ideas that the last chapter can stand alone. But I suspect that for most readers a prior acquaintance with at least the introduction will be necessary. Indeed, the very first paragraph of the introduction says that the latter will contain a number of 'indispensable definitions', which is curious in relation to the suggestion that it should come last. And the epithet 'indispensable' is no exaggeration. It may, however, be wise to follow the advice on procedure in so far as it entails reading the conclusions in a preliminary way before the central chapters, for the results of the many separate inquiries are presented in inevitably confusing detail.

The bulk of the research reported deals with classification rather than with serializing. This, we are told, is because classification is the more complex activity and the more difficult to understand. It is also surely the more important and so has been the more widely investigated in the past. But it is safe to say that there has been no previous developmental study of classification to approach this one in range. Some twenty different but closely connected studies are reported. They deal with the classification of geometrical forms and of objects; with 'free' classification and with classification controlled by specific requirements; with additive and multiplicative classification; with problems of inclusion and with the use of the quantifiers 'some'

and 'all'; with classes negatively defined; with the unit class and the null class; with the ability to anticipate a classificatory system in advance of the actual sorting of objects; with the ability to change criterion and reclassify; and finally with classification based on tactile-kinaesthetic perception. Some of these are studied in very simple and direct ways but in other cases there are procedures ingeniously devised to resolve special problems, e.g., when a weighing machine and boxes of different colours and sizes, as well as different weights, are used in order to study the child's understanding of 'all' and 'some' in a situation that is likely to seem to him to have real significance.

This particular test shows in very large measure a trend which almost all the tests show to greater or lesser extent: the trend away from reliance on the child's words and towards reliance on his acts which has become evident in Piaget's work over the years. It might not be too inadequate a summary of the book to say that it consists in an attempt to show that, in the absence of special inquiry, the child's ability to handle language may grossly mislead us as to his ability to handle classificatory systems. In this, as in much else, the authors come close to the position adopted by Goldstein—indeed the resemblances between their findings with young children and Goldstein's findings with aphasic patients are very striking. At what Piaget calls 'stage 1' (the familiar division into three stages appears once more)—the child produces 'collections *figurales*' where the tendency to build patterns and make 'thing-like' shapes is the distinguishing feature but where, as with Goldstein's patients, the patterns are obviously not planned in advance but arise through the accidents of the endeavour, the subjects being liable both to perseveration and undue fluidity of behaviour because they are at the mercy of the stimuli of the moment.

The inability to ignore spatial disposition, which Goldstein observes, has, indeed, a central role in Piaget's explanations of the development of classifying. His argument, in outline, is as follows. Classification is seen as having its origins in sensori-motor schemata, not primarily in language (which offers ready-made concepts certainly, but concepts that the child can accept only in so far as he can assimilate them), nor in perception except in so far as the latter is integrally bound up with the sensori-motor schemata. To understand the development of classification, then, precisely what one must understand is how classes differ from perceptual groupings and how, by means of 'operations' (which are in straight line of descent from the sensori-motor schemata) the child finally becomes capable of constructing the former and not merely the latter. The initial difficulty which the child faces in his attempts at classification is that, in perceptual aggregates, relations of resemblance between elements and spatial relations between them may both be ground for the decision that they belong together. Thus there is no regular correspondence between the qualities of elements and their association in a group. But the essential characteristic of a class is that there must be such a correspondence: that the qualities which define the class determine its membership. The qualities which define the class are known in logic as the 'intension', of the class. The members of the class constitute its 'extension'. Piaget and Inhelder maintain that in the young child these are not differentiated and hence cannot be co-ordinated. A large part of the child's difficulty in this matter is conceived as arising from—or at least manifesting itself in—his bondage to temporal succession, his inability to hold in mind and review objects are physically assembled, and hence physical dissociation destroys it. This means that the child can comprehend what Piaget and Inhelder call '*appartenance partitive*', or the relation of a part to a physical whole, but not 'inclusion' or the relation of classes in a hierarchy to one another. Thus even when, at stage 2, the 'collections' which the child forms become 'non-figural'—that is, are composed entirely of objects alike in some respect and are no longer patterning—the groupings remain for a time simply juxtaposed and do not belong in a true classificatory system: this even when it may seem on a first impression that the child is dividing and sub-dividing. To show how misleading this first impression may be, Piaget and Inhelder resort to the famous test, reported originally in *The Child's Conception of Number*, where a group of objects, *B*, is divided into two subgroups, *A* and *A*₁, and the subject is asked to say whether there are more *B*'s than *A*'s. The earlier results—that the young child maintains there are more *A*'s—are fully confirmed, and more extensive related studies of the child's understanding of 'all' and 'some' (one of which has already been mentioned) throw further light on the nature of the difficulty and form perhaps the most interesting and important part of the book.

These studies of 'all' and 'some' bring to a close the chapters which trace the main outlines

of the development of additive classification, but they are followed by a section where some special topics are considered, namely the child's ability to conceive of unit and null classes and to handle classes negatively defined. These researches are on the whole less impressive than the others, and there is one major objection to be made. This concerns the study of the empty or 'null' class—the class which has no members. The curious thing is that, although Piaget and Inhelder accept this usual definition (for they say that a null class is 'a class without objects'), they go on to describe a test where it is hard to see how any empty class can be held to be involved. The child is presented with a number of square, round and triangular cards, some having nothing on them at all. The latter, in other words are 'empty'. But a class of empty objects is plainly not an empty class. And yet it appears that the child's ability to conceive of an empty class is supposed to be tested by his readiness to make the dichotomous grouping 'cards with drawings: cards without drawings'. It is undoubtedly interesting to learn that up to the age of ten or eleven children tend to avoid this division, but the supposed dependence of the difficulty on that of the notion of the null class is quite obscure. In passing it should perhaps be remarked that Victoria Hazlett found children under the age of six able to make a distinction between cards with and without stars but admittedly the task she set them was slightly different from the one set by Piaget and Inhelder. In any case, whatever may be the arguments about the age at which the difficulty of the 'with:without' distinction can be overcome, the most reasonable view of its nature would seem to be that what is involved is difficulty in dealing with a class defined by the absence of a property—and formally this amounts simply to negative definition. 'Cards which have no drawings' and 'cards which are not round' both define negatively. Psychologically it is true, there may be a difference. Cards which are not round have other positive properties of shape, whereas cards which have no drawings are perhaps psychologically speaking, more truly defined by absence or blankness. But they certainly do not constitute a null or empty class.

The concept of 'the null class' is, of course, a highly sophisticated one. Strictly speaking, in an extensional logic there is only one null class, in the sense that no difference in membership can exist which could enable one empty class to be distinguished from another empty class. This is obviously a notion which we should expect children, not to say most adults, to find impossibly difficult. But even the simpler idea of a class with no members would seem to be impossible if—and for so long as—a class is for the child merely a physical collection, for presumably he cannot then conceive of a class so defined that there is nothing to collect. The difficulty of studying the early development of the notion of an empty class is that one must give the child physical objects to sort and yet, by definition, there can be no objects to represent a class which is empty. However, one might gain some information in the context of a study of multiplicative classification where things could be so contrived that an empty category would be necessary if a complete matrix was to be formed.

Piaget and Inhelder report studies of multiplicative classification but make no further mention of the null class. Their main concern is rather with the question of whether simultaneous classification by more than one criterion is indeed—as would seem likely on the face of things—more difficult than classification by one alone, and they arrive at the conclusion that this is not the case. The two are found to develop in close association, and if one is a little easier than the other then paradoxically it is multiplicative classification, or classification by more than one criterion, which seems to present least difficulty. This, the authors believe, is because the perceptual symmetries of the matrix layout are a help; and they comment on the relevance of their findings to tests like Raven's, which are based on the principle of matrix completion. But in the context of these tests can be solved either analytically or by global impression. But in the context of Piaget's 'operational' theories this suggestion becomes serious criticism.

The studies of multiplication having been completed, the rest of the book is concerned to demonstrate the theoretical importance of evidence of anticipation in a child's activities, as well as the existence and theoretical importance of a variety of relationships like the one between addition and multiplication which has just been mentioned. The authors find that classification by visual cues develops at about the same time as classification by tactile cues. Similarly with seriation, visual and tactile, simple and multiplicative: all are '*solidaires*' or integrally bound up with one another. The theoretical point is that this 'solidarity' is taken as evidence that the determinants of these activities are indeed the fundamental '*groupements*' which are postulated.

This emphasis on the importance, for the 'operational' theory, of the discovery that certain

abilities develop together is, of course, no new one. But we have had to be content until now with statements about ages and stages unsupported by any numerical data. For the first time these are offered here. The authors, having been, as they say, too often suspected of constructing theories on the basis of studies of 10 or 20 children, now tell us that they wrote this book after testing 2159. And they proceed to give numerous tables in which percentages of success at various ages are reported. This is of course a very welcome change of policy; but one cannot help wondering what now becomes of the arguments that were formerly advanced against the publishing of data of this kind. Piaget maintained in the past that one could not legitimately quantify so long as one retained the flexibility in testing that was necessary to discover what the child was thinking. Now he yields to critics without making it clear for what reasons he has decided that his own earlier arguments can be laid aside. Given, however, that quantification is now acknowledged to be possible, there is one surprising omission: no data concerning correlation are provided. We have already seen that the theories place great emphasis on the integrated nature of development. It is odd, then, to find instances where the same subjects figure in more than one table and yet not to be told fully and explicitly how the two sets of results relate to one another; for evidence of the relatedness of different abilities in the same children is obviously more effective support for conclusions about the integrated nature of development than evidence merely of a general tendency for the abilities to develop about the same time. However, regret that some information is withheld must not be allowed to overshadow the significance of the information which is provided. The giving of quantitative results adds greatly to the interest of what is undoubtedly a major study of a particularly important group of skills.

MARGARET DONALDSON

The Gifted Group at Mid-Life. By LEWIS M. TERMAN and MELITA H. OGDEN. Stanford University Press; London: Oxford University Press. 1959. Pp. xiii + 187. 36s.

This, the fifth volume of *Genetic Studies of Genius*, reports the results of a survey of the gifted group made in 1955. Most of the subjects, thirty-five years after the original inquiry, are approaching 50 years of age and are therefore fairly well established in a more or less permanent niche in society. The late Prof. Terman and his colleagues have achieved remarkable success in maintaining contact with the gifted group, and the fulness of the reports implies a very considerable amount of field and clerical work. The members of the group also appear to have been unusually co-operative.

The findings of this stage of the inquiry are not unexpected. Physical health is, if anything, above average; mental health is apparently average for the U.S.A. The high level of intelligence, on which the group was initially selected, remains and on a 'concept mastery' test the group retains its high intellectual status. Occupation-wise, the group has to a large extent justified its high intelligence. An interesting, but rather short, section deals with the children, of whom there are 2500, with a mean I.Q. of 133 on the Stanford Binet, the average I.Q. of the 'gifted' parents being about 150. There are also some grand-children.

The report reinforces earlier findings, which should dispel any lingering belief that gifted children die young, are physically weak or mentally unstable. The book represents a thorough and objective piece of reporting. The authors have difficulty in ascertaining national norms for comparison, e.g. in mental health, but are assisted greatly by the willingness of their subjects to answer detailed questionnaires; it is difficult to imagine a Scottish group, for example, being willing to answer seriously a question of which one item reads 'Alcohol is a serious problem. I am frequently drunk and attempts to stop drinking have been unsuccessful'.

This will be the last volume in which Prof. Terman will participate. He contributed a considerable part of the present volume, and though it is perhaps not the most interesting or distinguished of the series, it does show those qualities of thoroughness, objectivity and human sympathy which characterize the whole study.

JAMES MAXWELL

Manual of the Maudsley Personality Inventory. By H. J. EYSENCK. London: University Press. 1959. Manual 2s. 6d. net; Questionnaire 5d. each 4s. 3d. per dozen, 8s. 2d. per 25, 32s. per 100 including purchase tax; diagnostic keys 1s. each.

This Manual reports up-to-date information available on this Yes, No, ?-type questionnaire (*MPI*) comprising two scales of 24 items each, one purporting to measure neuroticism (*N*), the other introversion-extraversion (*I-E*). (The notation *I-E* is here substituted for Eysenck's *E* since the latter is inconvenient when referring to the introverted (*I*) or extraverted (*E*) end of the dimension.) The scales can be conveniently adapted to form an even shorter questionnaire (*SMPI*) comprising six items each for *N* and *I-E*, simply by utilizing the first page of the printed form only. The correlation of *MPI* with *SMPI* is 0.86 for *N* and 0.87 for *I-E*. The scoring is easily accomplished with the aid of templates using a simple numerical system.

The questionnaire, which owes a considerable debt to Guilford's original *R* and *C* scales, has undergone successive developments at Maudsley. Its present form is based on item and factor analysis of responses from a sample of 200 men and 200 women—but its standardization has since been supplemented by other groups ('dysthymics', hysterics, psychopaths, recidivist prisoners, psychosomatics, American students and a quota sample of the whole English population, the size of the sample for the latter being 1800).

The re-test reliability of the long scale is estimated as 0.83 for *N* and 0.81 for *I-E*; split-half reliabilities for the short scale are 0.80 for *N* and 0.72 for *I-E*. ($n = 2000$).

The two scales *N* and *I-E* intended to be orthogonal, have a low correlation (-0.15 for the *MPI*, and -0.05 for the *SMPI*) for normal samples—the correlation increases to the rather unsatisfactory dimension of -0.3 to -0.4 in neurotic groups. Eysenck assigns these anomalies to the non-linearity of the regression lines. The argument is supplemented by a graph of regression lines for 1200 normal subjects which does not show a serious state of affairs except *I*-scores below about 10. But the explanation as it stands is hardly sufficient to give rise to such large negative correlations between *N* and *I-E* in neurotic groups, whose mean scores lie well within the distributions given for the whole range of the population. This needs investigating in more detail. Meanwhile, when attempting to assess the effects of varying degrees of extraversion in any experiment, Eysenck suggests matching criterion groups of *I* and *E* for *N*.

A table containing the size of samples, mean scores and s.d.'s for the different standardization groups is given—but the reader is left to work out the significance of the differences. Unfortunately, the original sources of much of the data are omitted, or are still not available, so that the procedures whereby subjects were selected cannot be ascertained in detail.

Two methods of validation are presented: (i) comparison of the standardization groups on *N* and *I-E*; (ii) construct validity—i.e. a set of interlocking predictions forming a theory confirmed by experiment. In a strict sense, neither method can yet be said to have reached satisfactory standards in empirical confirmation. Our discussion must here partly traverse the boundary of the Manual, and consider also the references upon which the data in the Manual rests.

Let us first consider the *N* and *I-E* means and standard deviations placed graphically in their relative positions in the two-dimensional space formed by *N* (ordinate) and *I-E* (abscissa) dimensions. One might include ± 1 s.d. from each mean, extending the points by lines joined to form rectangles representing the expected scores within which 66% of the standardization groups are expected to fall. We then see (i) that there is a considerable overlap of distributions—i.e. the predictive power placing an individual definitely in one or other rectangle is not very great—but is greatest between dysthymics and normals; (ii) Moreover, it is immediately noticeable that the means of the groups fall about a line directed north-north-east. That is, while the abnormal groups tend to be considerably more neurotic than the normal English quota sample, there is little differentiation on the *I-E* dimension. Furthermore, what little differentiation there is (and it is reasonably considerable between normals and dysthymics) tends—apart from hospital psychopaths—to lie in only the *I* direction relative to normals. Put another way, unusual answers to some two to three items are sufficient to place an individual amongst the most extreme group of dysthymics on *I*, whereas some ten unusual answers are required to place an individual amongst dysthymics on *N*. At the very least there is some reason for attempting—in subsequent versions of the test—to stretch the *I-E* dimension somewhat. But the

reviewer is not certain whether the data cannot be taken as undermining one of Eysenck's basic tenets, which is not merely that hysterics are more extraverted than dysthymics but that they are also more extraverted than normals. In view of Hildebrand's similar findings with objective tests (*Brit. J. Psychol.* 1958, 49, 1-11), there is an increasing likelihood that the position of these various abnormal groups is a true feature of this *I-E* dimension, and not some distortion in its scale units—particularly at the *E* end. In fact, Eysenck hints at this situation in the Manual and elsewhere, without explicitly recognizing that it contravenes his and Jung's theoretical position. A factor to be taken into consideration, however, is that presumably this scale measures sociability rather than the other facets of extraversion emphasized by Jung—since *I-E* correlates highly with the 'social' scale and lowly with the 'thinking' and 'emotional' scale of the Minnesota *TSE*. Another feature worth noting is the limited range of items, many of which closely overlap in content. There might here be a tendency to sacrifice validity for reliability. Possibly a source of distortion on the *N*-scale, on the other hand, resides in the fact that in all items neurotic responses are scored in the affirmative (yes).

Space does not permit an appraisal of Eysenck's attempt to demonstrate 'construct validity' in his *Dynamics of Anxiety and Hysteria*, as claimed in the Manual. However worthy this attempt was, there has been a growing tide of criticism (e.g. by Storms and Signal, Vernon Hamilton, D. E. Broadbent, R. L. Reid, Taylor and Rechtsscaffer, Spivac and Levine) of the evidence presented for the various components of the theory in this monograph. Eysenck's adroit defences do not altogether dispel these criticisms. Consequently, the reader may have to exercise caution in accepting the construct validity as indeed valid.

Generally, Prof. Eysenck is to be congratulated on obtaining such an unusual amount of data on one personality test. It is to be hoped that a more detailed Manual will soon make its appearance; some of the original work is published in rather inaccessible foreign journals. In pressing the questionnaire as a research tool for obtaining criterion groups, it should, however, be hoped that the ease of using such a tool will not restrict the volume of research which ought to be applied directly to psychiatrically diagnosed neurotics. There is the danger that subjects selected on an *N* and *I-E* basis alone will unduly bias and filter the human material and cause much of importance to the clinician to be omitted. Conversely, it might not be altogether foolish to suggest that groups of subjects isolated out on the basis of *N* and *I-E* might, in addition, be subjected to routine psychiatric interview to preserve the 'feel' of the material and its possible limitations.

A. BURSILL

The Fifth Mental Measurements Yearbook. Ed. by O. K. BUROS. Highland Park, New Jersey: Gryphon Press. 1959. Pp. xxvii + 1292. \$22.50.

'Indispensable' is the word that comes most readily from one's pen in attempting to review a *Mental Measurements Yearbook*. The term Yearbook is of course strictly a misnomer, since this is, as the title tells us, only the fifth of the series inaugurated by *The Nineteen Thirty-eight Mental Measurements Yearbook of the School of Education, Rutgers University*. (Still earlier precursors were so much more modest in size and scope that they can be and evidently are by the author—regarded as preliminary ventures only.) In view of the enormous magnitude of the work it would be unrealistic to demand more frequent publication, although some things are bound to go out of date between successive Yearbooks, e.g. information regarding publishers, new or revised editions of tests.

While it seems inconceivable, to the confirmed Yearbook user, that any psychologist concerned with mental measurement in its broadest sense (educational assessment, projective techniques and questionnaires are all included) should not know what can be found in the Yearbook, it may be appropriate to explain that the Yearbook consists of two main sections. The first (extending in this edition to over 900 pages) surveys all tests in current use; bibliographical references covering all mentions in the literature are included, but it is disappointing to note that this particularly valuable feature is appearing for the last time. The second part (in this instance of about 200 pages) deals with books in all relevant fields; again the coverage is wide. Entries in *Psychological Abstracts* are noted, and the main content (as in the case of the Tests section) consists of entire or 'excerpted' reviews from a wide range of periodicals. Apart from the tenor of the reviews, no evaluation of the books or tests is offered. This may or may not be a good thing; the present reviewer is inclined to the former view.

British psychometry is unusually adequately covered—in fact there is little one could complain of in this respect, unless it were that the *Quarterly Journal of Experimental Psychology*, which has published relevant reviews during the period under survey, does not appear in the 'Periodicals Directory and Index'. It is pleasant, too, to note that the dedication of the book is 'To Cyril Burt'.

B. SEMEONOFF

A Textbook in Rorschach Test Diagnosis for Psychologists, Physicians and Teachers.

By EWALD BÖHM, translated by Anne G. Beck and S. J. Beck. New York and London: Grune and Stratton. 1958. Pp. xi + 322. \$7.75.

This book represents the most systematic and ambitious attempt yet published to produce a comprehensive guide to Rorschach diagnosis. It is systematic in that the author not only explicitly identifies himself with the psychoanalytic approach to personality study (as indeed most Rorschach practitioners do at least implicitly) but categorically states 'we do not believe that a Rorschach test can be exhaustively interpreted without knowledge of psychoanalytic theory.' This remark occurs near the beginning of a long chapter (of over 50 pages) which presents the first and most detailed of a series of discussions of psychopathological conditions in which the Rorschach test response is used to illumine theory, rather than vice versa, as is more usual in Rorschach literature. It is ambitious in so far as guidance—often avowedly 'rule of thumb'—is given for much more refined diagnosis than Rorschach writers usually offer. Nevertheless, in a footnote (pp. 240–1) the author writes: 'How much or how little is required for a diagnosis is a matter of judgement, and depends on the correct evaluation of the psychological meanings of the various factors.' He then goes on to warn the reader against mechanical counting of 'signs'. The same footnote ends with the announcement that a 'summary in tabular form of all syndromes contained in this book is reserved for a later publication.'

A further valuable and unusual feature of the book is its frequent reference to the work of European psychiatrists and Rorschach writers, notably Binder, Zulliger, Oberholzer and Loosli-Usterli, whose work is relatively unfamiliar, at any rate in detail, to British or American psychologists. Unfortunately the continental bias makes itself felt by the presence in the text of large numbers of pseudo-technical terms, some based on false etymology and nearly all open to condemnation by the authors of a recent psychological dictionary as examples of bogus erudition. Thus such words abound as amphithymia (and also *ambithymic*), ixothymia (and ixophrenia), schizaffinic, pathotropic, encephalopathy, pprofschizophrenia, sophropsychic (thus in the text; in the index—more correctly?—*sopropsychic*). One may perhaps be forgiven if one confesses that the vulgarity 'baffling with science' comes to mind, and if one points out, by way of retaliation, that four mistakes appear in the printing of the six-letter Greek word from which 'ixothymia' is derived.

All this makes for heavy going, and—again regrettably—the translation itself has a marked continental flavour. Furthermore, although the book is intended for the beginner as well as for the specialist, some of the advice given to the learner seems to the reviewer to be positively harmful, e.g. the recommendation that locations should be noted in words rather than by the use of the location chart. The conclusion one seems to come to is that this is not a book for the beginner, nor for those only partially convinced of the value of the Rorschach method. For the advanced practitioner, on the other hand, it opens out new vistas.

B. SEMEONOFF

The Szondi Test in Diagnosis. Prognosis and Treatment. By L. SZONDI, V. MOSER and M. W. WEBB. Philadelphia: Lippincott; London: Pitman, 1959. Pp. xv + 309. 96s.

The authors have attempted to compress into one book the four works of Szondi on his personality test and its application to diagnosis, prognosis and treatment as well as further work upon it. Since the early days, this test has grown into a psychology and a method of analysis now called Schicksal psychology and Schicksal analysis, the word meaning 'fate' or 'destiny'. Considerable claims are made for the method, for instance, that it is the 'Rosetta stone of

psychology'. As for the theory, it purports to combine heredobiology and psychoanalysis in a unique manner.

The authors have made their task more difficult, and the book less readable, by their anxiety to prove the superiority of their theory and method to others, and by trying to confound those psychologists who rejoice in making nonsense of tests and theories which rely on the concept of a dynamic unconscious. The somewhat messianic attitude goes with the scolding meted out to the unconverted and seems more likely to put people's backs up than to induce them to take an interest in the material presented. Unfortunately, clarity is further impeded by the language used, which is mainly in the Teutonic style, heavy, often florid and sometimes obscure. Reading is made difficult by the Szondiian terminology which, in his own works, is explained at length as he builds up his concepts, but in this book is piled upon the reader in a more or less arbitrary manner.

The book is in three parts. The first discusses the rationale and validation of the test and describes the procedure involved in the testing of individuals both singly and in groups. It ends with a chapter by Szondi himself on what might be called his philosophy, which would perhaps have been better put first as it serves to orientate the reader towards what follows. His central concept is of the *familial unconscious*, inherited through the genes, which influences the individual in his choice of love objects, ideals, occupation, illnesses and form of death. Thus, he calls Schick-ego and the intellect so that if the person becomes aware of the morbid ancestors of whom he is a carrier or conductor, their influence may be modified or transformed. Abstracting from the genetic basis of the theory, which seems superfluous from the point of view of therapy, it differs little from that of the internalized object of the Neo-Freudians and from the ancestral images of Jung. Szondi's contribution is that his test specifies them so that one has a clear picture of the images the person harbours in the unconscious, of the areas they afflict and of his ego defences against them. Further, changes can be predicted, and possibilities for development too.

The second part of the book is on the clinical application of the Szondi test. Studies of a hermaphrodite and of an epileptoid murderer are given at length, and there are two statistical studies. One is on factorial variability and the other on the theory of probability in relation to the test.

Part three is written by Szondi on the therapeutics of Schicksal analysis. He goes more fully into his theory and describes his method of treatment and its aim. Starting with what he terms a formal analysis of repressed material, he proceeds to his 'Hammer-blow' technique by which means he forces the patient to abreact the latent morbid ancestors in full consciousness. This is followed by 'ego analysis'. He deals with various objections raised to his technique. The reviewer's objection is that he starts with a plan of what he thinks is good for the patient and then hammers it into him. His aim is a benevolent one, that of socialization and 'humanization' test. It may become more than a guide to the appropriate therapeutic approach—it becomes the blue-print from which no divergence is possible.

MARY WILLIAMS

Hérédité et Fréquence des Dyschromatopsies. By R. KHERUMIAN and R. W. PICKFORD.
Paris: Vigot Frères. 1959. Pp. v + 109. No price given.

The introduction to this small book disarms some criticism by stating the limited aim, which is a survey of the results of studies of the frequency of colour vision defect in various populations of the world.

The first part of the book, 44 pages out of a total of 106, may be regarded as giving the background necessary to understand such a survey. There are short chapters on the definition and classification of kinds of colour vision defect, including an historical survey of the reports of colour blindness published before and after John Dalton's account of his own case. The classification adopted is that used by Pickford in his earlier writings. A short chapter deals with the colour vision of animals and the theories of colour vision. The next chapter describes clearly what is known about the inheritance of colour vision defects, using mainly the pedigrees published by Julia Bell in 1926, and goes on to discuss the frequency of various kinds of colour vision defect if the assumptions about its genetic mechanism are correct. The 'problem of twins'

(identical twins especially) and 'aberrant pedigrees' are discussed briefly, including one or two recent pedigrees, and it is suggested that in order to deal with the problem of the genetics of colour blindness we need a combined operation by ophthalmologists, physicians, biochemists and geneticists.

Other problems are dealt with in tantalizingly short accounts; e.g. the colour blindness found in what are called 'inter-sex groups' (the Turner and Klinefelter syndromes), other sex-linked defects not apparently related to colour blindness, acquired colour blindness and colour vision and age.

The second part of the book reports the results of large-scale testing of University students in Paris carried out by Kkerumian and his colleagues over a period of years, only some of which has previously been reported. Besides using a large number of pseudo-isochromatic plate tests, Kherumian noted the place of birth of the subjects and of their parents, and was thus able to compare the frequency of colour blindness in the various Departments of France, and in the French 'associated states' and other foreign countries. The total number tested was 9151 men and 8945 women. Although the frequencies in the Departments varied considerably (e.g. 4.12% in the Moselle Vosges Departments and 16.3% in Gard Hérault Departments), these differences turned out not to be significant. These results are contrasted with those reported by Vernon and Straker, in *Nature* in 1943, which are also reported here in a rather fuller form than was possible in 1943. Some racial differences are found; the frequency of colour blindness is about 8% for Metropolitan France, and most of Europe and America, but only 6% for the students from Algeria, Morocco, Tunisia, etc. It is suggested, further, that in the latter group, the proportion of protanopes is much higher than it is in the other colour blind groups. However, this statement is based on the comparatively small proportion of all the colour-blind subjects who could be clearly classified as protanope or deutanope on the Plate tests, and may be thought not to be too well established.

In the penultimate chapter, the authors survey briefly the literature on the frequency of colour blindness in different parts of the world, dismiss many of them as being quite inadequate, including, alas, the much quoted Rivers's study of the Todas, and then make an annotated list of the studies which survive their critical evaluation. The final chapter summarizes the evidence about the frequency of colour blindness throughout the world, and points out that the figure of about 8% for the white European, American and Australian populations is higher than that for any other racial group; it is slightly less for the Asiatic people, and even less for Negro populations. The differences cannot be related yet, to any precise anthropological criterion of race, and are in quite the reverse order, they say, to that which one would expect on any evolutionary theory of the development of colour vision, or of races.

Lest human interest should be entirely lacking, they have included photographs, test records and descriptions of two pairs of identical male twins, both colour blind.

The bibliography, in which one should include the annotated list of Surveys, is fairly extensive.

In general the merits of the book are difficult to assess. Much of the material presented is already familiar to those interested in colour vision and colour blindness, yet the accounts given are too brief to be useful to the beginner. They are perhaps sufficiently stimulating to send the interested reader to the original papers if he has not already seen them.

AGNES CRAWFORD

Ego Structure in Paranoid Schizophrenia. By LOUISE J. ZUCKER. Illinois: Thomas; Oxford: Blackwell. 1958. Pp. x + 186. 42s.

This is a report of an experimental study, by a clinical psychologist, of disturbance of ego function in paranoid schizophrenia. In this study, three projective tests were used—the Rorschach, Mosaic, and Figure Drawing. The subjects were two groups of paranoid schizophrenic patients, thirty in each group, one group being seriously ill patients (the 'hospitalized' group), the other group was of paranoid schizophrenics who were less incapacitated by their illness and who were being treated on an ambulatory basis.

Dr Zucker aimed, first, to devise a technique of identifying from the test responses, the degree of impairment of ego structure in each patient, and then to compare the two groups, the hypothesis being that the hospitalized group would show a greater degree of dysfunction than the ambulatory group. She begins by defining five 'signs' in the test which, on theoretical grounds, she takes to indicate ego dysfunction, and in comparing the two groups with respect to the frequency of the appearance of these 'signs', she finds a significantly higher number in the hospitalized patients as compared with the ambulatory group. In a second phase of the experiment, she reports that the same test method differentiates significantly between the 'ambulatory' patient who can respond to treatment on an out-patient basis, and the clinically similar ambulatory patient who later will break down and who will require hospitalization. This is a very important prognostic finding which will be of great interest to all clinical psychologists and psychiatrists concerned with diagnosis and prognosis in schizophrenia.

The book is well written and the research method clearly described in detail, so that the experiment should easily be repeatable. Her work opens up many avenues for further exploration. If her results are confirmed, she will have made a significant contribution to clinical psychology and psychiatry.

J. L. BOREHAM

Psychology and Human Performance: An Introduction to Psychology. By ROBERT M. GAGNE and EDWIN A. FLEISHMAN. New York: Holt. 1959. Pp. xvi + 494. \$7.25.

The authors of this introductory text have had a number of definite aims and presumably particular groups of students in mind in writing their book. The last third of the book is about the industrial application of physiology and deals with jobs and personnel selection, training, and human engineering, and the earlier part is essentially prefatory to this. The emphasis is on the capacities and performance of the human being, and wherever possible illustrations are taken from the vocational field. The cybernetic-information theory model is used as the framework upon which the text is constructed. The major divisions of the human behaviour system are concerned with the functions of receiving stimulation, transmitting signals, acting, and energizing. The human being possesses certain basic abilities which are added to by learning. These abilities are considered under the major heads of discrimination and identification, motor skills, concepts and thinking. A chapter on Social Behaviour is concerned mainly with the industrial aspects of group behaviour.

Within its chosen and rather narrow limits the book is a good text, clearly written and well illustrated, particularly with numerous groups. It should be specially useful as a first year introduction to psychology for those who intend to specialize later in the industrial human-engineering field. Its limitations, however, would preclude its use as a general first-year text in most university classes, as it omits nearly all reference to a good many topics upon which the general student should be informed, for example, development, childhood, emotion, personality, adjustment, animal behaviour, etc. The book is also excessively parochial in its references. There are fewer than a dozen in the whole field of psychology to non-American work—none, for example, to Freud—but no less than thirty-five references to the authors' own studies.

L. S. HEARNshaw

The Problems of Perception. By R. J. HIRST. London: Allen & Unwin, 1959. Pp. 330. 30s.

Bertrand Russell recently remarked that, for the new philosophy, 'it is regarded as a sin to think seriously about the problem of perception'. In writing this book Mr Hirst, has on this reckoning, committed a very grievous sin indeed. For, instead of dismissing in a few glib phrases, as is the custom, everything that has been said on the topic, in an effort to persuade us that there are no problems to solve—we have all just been the victims of a misunderstanding—he has examined carefully and respectfully a variety of philosophical theses before eventually offering us his own considered solution. What is perhaps even more unusual and refreshing nowadays, is to find a philosopher who can discuss perception with a proper understanding of recent experimental work. His chapter on 'The Psychology of Perceptual Consciousness' is helpful and makes some valuable terminological suggestions.

After taking us on a tour during which we are shown the perplexities of sense-data theories, the implausibility of phenomenalism, and the sheer evasiveness of the linguistic approach, he finally introduces us to the only two types of solution that can seriously compete for our allegiance: a dualistic, causal or 'Representational' theory on the one hand and a monistic, 'Double-Aspect' theory on the other. Despite a sidelong glance at E.S.P. the author comes out firmly in favour of the latter. What he calls his 'Identity Hypothesis' asks us to treat both the brain activity and the conscious perceptual experience as simply the outer and inner *aspects*, respectively, of one and the same physical event. The distinction is merely in mode of access: brain activity being (in principle) publicly observable, perceptual experience only privately introspectible.

Such a view (which on my reading of it is essentially the same as that now held by Russell) has obvious advantages: it is economical, it minimizes mystery, and it leaves intact the deterministic presuppositions of the natural sciences. Nevertheless, Hirst lays himself open to certain legitimate criticisms. For, on a mode-of-access theory, who or what is it that does the introspecting? Moreover, how exactly does such introspection differ from that 'awareness of one's one sense-data' which Hirst regards as introducing a vicious reduplication into the Representational account of perceiving? It is true the author makes a valiant effort to answer both these questions, the former on p. 302, the latter on p. 293, but the reader will have to decide whether he finds his answers entirely satisfactory.

Hirst evidently shares the current malaise about mental entities and this, it seems to me, sometimes drives him into unnecessarily strained positions. Thus he claims with obvious pride to have purged the Representational theory of its 'extra mental realm'. Yet his own analysis compels him to recognize the act/object character of perceptual consciousness and to speak freely of the *contents* of such consciousness. By the time he comes to discuss the nature of sensible properties (p. 314) we may be forgiven for wondering whether he has ever really broken away from the Representational theory as opposed to merely trying to make it more acceptable.

These few remarks, however, must not be understood as implying any lack of appreciation for what I regard as a very considerable achievement. I have no hesitation in recommending it to any psychologist interested in the philosophical implications of perception.

JOHN BELOFF

OTHER PUBLICATIONS RECEIVED

- Rorschach et Electroencéphalogramme Chez L'Enfant Épileptique.* By Z. HELMAN. Paris: Presses Universitaires de France. 1959. Pp. 403. 2000 F.
- La Sélection Professionnelle.* By S. PACAUD. Paris: Presses Universitaires de France. 1959. Pp. 178. 7 00 NF.
- Estimation of Time.* By M. FRANKENHAEUSER. Stockholm: Almqvist and Wiksell. 1959. Pp. 135. Sw. kr. 16.
- The Campden Wonder.* Ed. by Sir GEORGE CLARK. Oxford University Press. 1959. Pp. 155. 18s.
- Papers on the Teaching of Personality Development. The Sociological Review Monograph No. 2.* Ed. by PAUL HALMOS. Keele: University of North Staffordshire. 1959. Pp. 149. 12s.
- A Preface to Scientific Psychology.* By RAYMOND J. MCCALL. Milwaukee: The Bruce Publishing Company. 1959. Pp. 74. \$1.00.
- Suggestibility: a Factorial and Experimental Analysis.* By K-G. STUKÁT. Stockholm: Almqvist and Wiksell. 1958. Pp. 248. Sw. kr. 25.
- Personality Patterns of Psychiatrists.* By R. R. HOLT and L. LUBORSKY. London: Imago. 1959. Pp. 386. 42s.
- PAT Interpretation.* By S. S. TOMKINS and J. B. MINER. New York: Springer. 1959. Pp. 184. \$5.50.
- Automatic Teaching.* Ed. by E. GALANTER. New York: Wiley. London: Chapman and Hall. 1959. Pp. 198. 26s.
- Reflections on the Human Venture.* By H. CANTRIL and C. H. BUMSTEAD. New York University Press. 1960. Pp. 344. \$6.50.
- In Place of Parents: A Study of Foster Care.* By G. TRASLER. London: Routledge and Kegan Paul. 1960. Pp. 248. 25s.
- Francis Bacon.* By J. G. CROWTHER. London: The Cresset Press. 1960. Pp. 362. 35s.
- L'Intelligence et Le Caractère.* By R. MAISTRIAUX. Paris: Presses Universitaires de France. 1959. Pp. 352. 15 00 NF.
- The Mother-Child Interaction in Psychosomatic Disorders.* By A. M. GARNER and C. WENAR. Urbana: University of Illinois Press. 1959. Pp. 290. \$6.00.
- Anthropology and Ethics.* By M. EDEL and A. EDEL. Springfield. Illinois: Thomas. 1959. Pp. 250.
- Malformations Congénitales du Cerveau.* By G. HEUYER, M. FELD, J. GRUNER and others. Paris: Masson. 1959. Pp. 450. No price given.
- Teaching Comprehensive Medical Care.* By K. R. HAMMOND and F. KERN. Harvard University Press. London: Oxford University Press. 1960. Pp. 642. 80s.
- Freud: The Mind of a Moralist.* By P. RIEFF. London: Gollancz. 1960. Pp. 397. 30s.
- Understanding Human Society.* By W. GOLDSCHMIDT. London: Routledge and Kegan Paul. 1960. Pp. 253. 21s.
- Outline of Human Genetics.* By L. S. PENROSE. London: Heinemann. 1959. Pp. 146. 12s. 6d.
- Design for a Brain (2nd ed.).* By W. ROSS ASHBY. London: Chapman and Hall. 1960. Pp. 286. 42s.
- Personality Change and Development as Measured by the Projective Techniques.* By M. HARROWER. New York and London: Grune and Stratton. 1958. Pp. 383. \$10.00.
- Institutional Neurosis.* By R. BARTON. Bristol: John Wright and Sons. 1959. Pp. 56. 8s. 6d. post 5d.
- Morality and the Homosexual.* By M. J. BUCKLEY. London: Sands. 1959. Pp. 214. 16s.
- The Pane of Glass.* By J. B. MARTIN. London: Gollancz. 1960. Pp. 397. 30s.
- Heilpädagogische Psychologie.* By P. MOOR. Bern and Stuttgart: Huber. 1960. Pp. 326. DM 32.80.
- Practice and Theory of Probation and Parole.* By D. DRESSLER. Columbia University Press. London: Oxford University Press. 1960. Pp. 252. 48s.

INHIBITION: A SYMPOSIUM

At the Cambridge Conference of the British Psychological Society a Symposium on 'Inhibition' took place on 6 April 1959. The individual contributions are now available for publication having been revised by their authors in the light of the symposium as a whole.

The first of these papers which follow comprises an introduction which states what the authors, Pilkington and McKellar, consider to be some of the more profitable uses of the term 'inhibition' in psychology. The second paper, by Stretch, deals with ideas and researches relating to inhibition within the evolutionary system as a whole. Kendrick's contribution is an account of experiments on the possibilities of establishing inhibition through reinforcement, and a defence of his theoretical interpretation of these experiments. In the fourth paper Gwynne Jones deals with the work of Eysenck and his colleagues in their application of Hullian and Pavlovian concepts of inhibition to the study of individual differences. The concluding paper by Reid presents a critical evaluation of the use made of 'inhibition' by Pavlov, Hull and Eysenck.

A bibliography of references cited by the various contributors concludes the symposium.

PETER MCKELLAR
University of Sheffield

I. INHIBITION AS A CONCEPT IN PSYCHOLOGY

G. W. PILKINGTON AND PETER MCKELLAR

Department of Psychology, University of Sheffield

One way to begin this Symposium is to look at what the dictionary has to say about the subject under discussion. In the larger *Oxford English Dictionary* 'inhibition' is defined as 'the action of preventing, hindering or checking'.

Both the larger and smaller editions of the dictionary make reference to an article by Brunton published in *Nature* in 1883. Brunton wrote: 'By inhibition we mean the arrest of the functions of a structure or organ by the actions upon it of another, while its power to execute these functions is still retained and can be manifested as soon as the restraining power is removed.'

Thus, according to the dictionary there is implied by the use of the word the notion of being checked or restrained by a process of an opposing kind.* This same notion of restraint is prominent in the three columns of definitions of 'inhibition' which appear in the dictionary of psychological terms by English & English (1958). Dictionary definitions of 'inhibition', whether general or technical, provide then a starting-point for this symposium. But it is only a starting-point. If we turn from dictionaries to text-books it becomes plain that 'inhibition' is a very widely used word in psychology; very few such text-books fail to include it in their index. It is equally apparent that uses of the term vary enormously, and this is probably related to the fact that the word often appears qualified by some adjective: for example we encounter 'conditioned inhibition', 'proactive inhibition' and so on. Uses of the term vary from the simple and descriptive to what some of us in this symposium may regard as the excess of 'physiologizing' to which many theorists in contemporary psychology seem so strangely addicted. And, of course, over a whole family of uses there looms the dark, and bearded, shadow of Ivan P. Pavlov.

It will be part of our aim in introducing this symposium to draw attention to the extremely wide *variety of uses* the term 'inhibition' has in psychology. Since bias may express itself implicitly if not made explicit, it is desirable that we define our own standpoint, which is twofold. First, we wish to stress the various different fields of theory to which some concept like 'inhibition' seems appropriate and useful; more specifically we shall argue that concepts and principles deriving from Pavlov and Hull—despite their undoubted importance—should not wholly dominate the symposium. Secondly, it seems to us that, while theories in psychology come and go, the *phenomena of behaviour and experience* will go on for quite a long time yet, and we do need terms to describe them. Thus, much of our emphasis will be on the phenomenological as well as the theoretical uses of the term 'inhibition'.

* This is the physiological use of the term. According to Fulton (1949, p. 93) its application to the C.N.S. relates to the work of the brothers Weber who, in 1845, showed that stimulation of the peripheral end of the vagus nerve produces temporary cessation of the heart beat.

The first of these uses with which we shall be concerned is when 'inhibition' is used descriptively to denote some established phenomenon of behaviour. Illustration may be taken from a story reported by Hilgard (1953). It relates to a former President of Stanford University, who happened to be an expert on fish. As President, he diligently and dutifully set about the task of learning the names of his students. But the worst happened. Every time he learned the name of a student he found he forgot the name of a fish! And so it was, we are told, that, under the influence of the inscrutable processes of retroactive inhibition, he gave up learning the names of the students. The realities of the phenomena of associative inhibition must be familiar, probably in a context of card sorting, to nearly every psychologist. Even if he is to-day unable, without some effort of recall, to remember which is 'pro-' and which is 'retroactive' inhibition, once he has thought about it, the meanings of these terms are fairly clear and precise. Whatever this symposium decides to do about 'inhibition' as a concept in psychology we should like to put in a word for this usage as a descriptive term—especially along with a qualifying adjective like 'associative'. Hilgard elsewhere (1948, p. 113) has ventured the prediction that, as psychology develops, its text-books will grow more and more like Hull's *Principles of Behaviour*. While this doubtless desirable process is going on, psychologists are likely to want to go on talking to one another. In this interim, if they can be encouraged to use qualifying adjectives with otherwise only vaguely descriptive terms like 'inhibition' some measure of precision will be added to their communications.

A second—and by now a classical—use of the term 'inhibition' is to be found in the work of Pavlov (e.g. Pavlov, 1927, 1955); we shall not however intrude upon later parts of the symposium by analysing these Pavlovian uses in detail. Nevertheless, it may be noted that our argument that the term should, where possible, be used along with some qualifying adjective receives strong support from Pavlov. He distinguishes a number of different forms of inhibition: internal inhibition, external inhibition, conditioned inhibition, differential inhibition, disinhibition and so on. One interesting point concerns internal inhibition which is, oddly enough, often spoken of as 'extinction'. The use of *this* word is surely paradoxical since it seems to imply that the CR has disappeared rather than that it is merely in a state of inhibition. This strange equation of 'extinction' (i.e. disappearance) with 'inhibition' (i.e. being really there but restrained from appearance) is now very common. It is probably too late to do anything about altering it, even though, as there are operational tests for showing that an extinguished response is, in a sense, 'really there', the term 'internal inhibition' would seem descriptively more appropriate.

A third, and also very well-known use of 'inhibition' derives from the work of Hull (e.g. 1943, 1951, 1952), sometimes via South London! Again we find 'inhibition' used along with an adjective, and essentially Hull distinguishes three types. First, there is *reactive inhibition* or I_R which is, in some respects, similar to Pavlov's internal inhibition. It is a kind of negative drive, produced whenever a response occurs, and acting as a barrier to the repetition of that response. In other words, reactive inhibition directly inhibits reaction potential (sE_R). It dissipates spontaneously with time, but has the important property (since it is a negative drive) of reinforcing any activity associated with its reduction. Next in the Hullian system is *conditioned*

inhibition (sI_R). This is perhaps best thought of as a habit of ceasing to respond* in a particular way in a particular situation. This ceasing to respond is initially brought about by a succession of increments of I_R , which totally, though only temporarily, inhibits sE_R . Hull's notion is that, when such stoppage of response has occurred a certain number of times, the habit of ceasing to respond in a particular manner in a particular situation is built up because of the reinforcement provided by the dissipation in time of the negative I_R . It is important to keep in mind that, as a habit—a conditioned disposition to cease to respond— sI_R does not dissipate spontaneously with time. Clearly there are difficulties in the way of conceiving of *ceasing to respond* as nevertheless 'really' being a sort of response that can be learned; and it is even harder to see how such a habit of ceasing to respond could possibly be subject to experimental extinction. But these difficulties will be dealt with more fully by later contributors. Finally, there is *aggregate inhibitory potential* or $\dot{I}_R$, and this Hull defines as simply the sum of reactive inhibition with conditioned inhibition. Again, there have been objections to this concept, mainly on the grounds that it involves the addition of a drive construct to a habit construct, whereas these should surely be multiplied together. These objections will also be debated in some detail later in this symposium.

From Pavlov has stemmed the work of Hull. However influential Hullian thought has been it is by no means the only attempt which has been made to construct a theory of learning in which inhibition plays an important part. A rather different development is represented by Harlow's recent and ingenious attempt to explain learning solely in terms of one process of inhibition rather than in terms of opposing processes of excitation and inhibition (Harlow & Hicks, 1957; Harlow, 1959). Harlow presents data, chiefly deriving from his experiments on the formation of learning sets, and on the elimination of error factors (E.F.'s) which, he claims, suggest that all learning may best be understood as the inhibition of responses rather than as the formation of new associations. He writes: 'the development of the thesis that learning involves nothing other than the elimination of responses and response tendencies inappropriate to a particular learning situation leads us to question the commonly held concept that learning is the resultant of two processes, excitation which follows reward, and inhibition, generated when no reward is received' (Harlow, 1959, p. 526). Clearly, such a uni-process, non-reinforcement theory of learning contrasts sharply with Hull's duo-process, reinforcement theory. Some psychologists may regard Harlow's claim that 'learning is *nothing but* (italics ours) suppression or inhibition of E.F.'s' as openly paradoxical—or, at least, as linguistically decidedly odd. But whatever the case, in Harlow's work the concept of 'inhibition' plays the key role—and in a modern theory of learning which may well assume much importance in future years.

A fifth use of 'inhibition' is to be found in much of the work of Eysenck and his colleagues. Eysenck's deployment of the concept of inhibition is probably the most self-conscious effort we have in modern psychology to make an extensive *theoretical*

* Originally we described sI_R as an acquired habit of *not responding* in a particular situation. We are indebted to a later contributor to the symposium, Reid, for distinguishing between such a habit of not responding, and a habit of ceasing to respond. Reid has pointed out that Hull used sI_R in the latter sense, and we have amended our paper accordingly.

use of the notion. At the same time he makes extensive uses—some may feel inclined to say ‘misuses’, if not ‘abuses’—of both Pavlovian and Hullian notions of inhibition. As used by Eysenck the concept of inhibition becomes a central one in psychological theory. It is used to explain and predict a very wide range of behavioural phenomena extending from the molecular (in physiological psychology) to the molar (in personality and social psychology). The details of this development of the concept of inhibition are to be found particularly in *The Dynamics of Anxiety and Hysteria* (Eysenck, 1957), and will be discussed later in this symposium. Here it is sufficient to say that a difference in inhibition is postulated as between introverts and extraverts: the introverts are held to build up inhibition slowly (hence to condition quickly and extinguish slowly), and the extraverts are held to build up inhibition quickly (hence to condition slowly and extinguish quickly). Eysenck contends that these differences of inhibition between the two groups have been confirmed experimentally, and that many deductions from his postulates which relate inhibition to personality variations (e.g. reminiscence and work decrement phenomena) have been similarly verified. It may be noted that Eysenck’s apparently so *theoretically* fertile concept of inhibition differs markedly from one important everyday use of the term ‘inhibition’. In common usage it is the introverted, socially diffident, and emotionally held-in sort of person who is described as ‘inhibited’, while the sociable, outgoing extravert is described as ‘uninhibited’. For Eysenck the position is exactly reversed: behavioural inhibition in the sense just discussed is considered to be inversely related to inhibition. It may also be noted that Eysenck (1957, pp. 112–13) takes Pavlov to task for extending these everyday uses of ‘inhibited’ and ‘uninhibited’ into his psycho-physiological theorizing.

A sixth, and rather different conception of ‘inhibition’ derives from the thought of Herbart. Herbart was at the time preoccupied with the problem of what would to-day be called ‘span of apprehension’ (or ‘narrowness of consciousness’), and was one of the first to explore the notion that mental processes in active opposition to one another may compete for a place in consciousness. This concept subsequently played an important part in Freud’s thinking. Freud’s theory of repression (e.g. Freud, 1938) implies that forgetting is not merely a matter of the fading of memory traces; on the contrary forgetting is sometimes (? or always) an active blocking of one set of emotionally toned ideas by another. Writing from a Freudian standpoint, and as an historian of psychology, Flugel (1955, p. 8) goes so far as to suggest that it is now generally agreed that repression ‘corresponds to inhibition, which is found at all levels of the nervous system’.* Psychologists who are outside the psychoanalytic camp might well consider more fully these efforts—so characteristic of Flugel’s own work—to relate analytic concepts to more general psychological and physiological principles. Another good example of an attempt to relate experimental and psychoanalytic work is to be found in the work of Dollard & Miller (1950) who examine several different types of ‘repression’ in terms of their relation to ‘inhibition’. Like Flugel, these authors equate repression with inhibition, and an illustration of their attempted reconciliation of psychoanalysis with learning theory may be given. They argue that

* Flugel goes on to suggest that relations between ‘repression’, as understood by Freud, and Sherrington’s ‘antagonistic reflexes’ merit fuller examination.

in war neuroses there develops a tendency to stop talking and thinking about the experiences responsible, and that fear strongly reinforces this response. More generally, if a person tends to stop thinking about certain facts he is less likely to see connexions between them; this is seen by Dollard and Miller as the entrance chamber to the process of repression. The attempts of learning theorists to understand psychoanalytic thought in terms of their own concepts is in itself a development of major interest in view of the historical connexions between these concepts and the work of Pavlov and Hull, already mentioned.

Further reference may be made to more strictly Freudian thought, in which 'inhibition' is by no means confined to the problems of remembering and 'repression'. For example, Freud's theory of laughter involves a rather similar concept. He suggests that inhibitory processes operate to prevent certain expressions of aggression or sexual allusion; these are relaxed in the socially permitted form of wit. Thus, in wit, there is a release of the affect itself together with a release of the inhibitory energy (which he later called anti-cathexis) involved in *maintaining* the taboo. Rickman (1957, p. 253) suggests that this conception of the accompanying release of inhibition in wit is helpful to the understanding of the manic-depressive psychoses: 'in depression inhibition is increased, in mania it is diminished'. Of the manic individual, he writes: 'his pleasure is similar to the pleasure of wit—the release of inhibition'. Freud himself conceives of the depressed person as emotionally preoccupied with something; when mania ensues the mental energy involved in this preoccupation is released. As Freud puts it in a picturesque phrase: 'the manic subject plainly demonstrates his liberation from the object which was the cause of his suffering, by seeking like a ravenously hungry man for new object cathexes' (Freud, 1917, p. 255). This concept of expenditure of cathexis and anti-cathexis as involved in emotional preoccupation, followed by subsequent release of the inhibition necessary to the maintenance of this preoccupation, would seem to be at least descriptively appropriate to the manic-depressive cycle.

Rather different from these psychoanalytic concepts of inhibition is the use made of the notion in physiological psychology. The most important of the non-Pavlovian uses derives from the by now classical work of Sherrington (1906) and of Hughlings Jackson (1931). The term 'inhibition' is employed to denote a reduction in, or prevention of, response, as a result of stimulation. The notion assumes especial importance when used to explain various aspects of reflex or other co-ordinated behaviour in which the successful occurrence of one response entails the inhibition of other antagonistic or incompatible responses. Indeed there is much to suggest that inhibition is a fundamental aspect of the activity of the C.N.S.; thus Eccles (cited Ritchie Russell, 1959, p. 123) considers that 'the cells of the C.N.S. may consist of two basic types—those which excite and those which inhibit'. Certainly 'inhibition' with various qualifying adjectives added to specify the type or locus is a very commonly used term; thus we have: 'reciprocal inhibition', 'successive inhibition', 'direct inhibition', 'extrapyramidal inhibition', and so on. The meanings of such terms are clearly defined, and the use of a qualifying adjective is in accord with our recommendation. In this connexion mention must be made of the central place which the concept of 'reciprocal inhibition' holds in the work and thought of Wolpe (1958), specifically in relation to therapy. Wolpe argues for a more direct approach

to the behaviour disorders than the traditional Freudian one, and contends that they are best conceived of in terms of the principles which underly habit formation, and their related neurology. Ordinary methods of treatment involving internal inhibition or extinction as studied by such workers as N. E. Miller are slow and laborious. Instead of employing extinction Wolpe advocates a process of setting up antagonistic habits, in short using reciprocal inhibition; this is interpreted both physiologically and in terms of the psychology of learning. The results of treatment by reciprocal inhibition on a group of 210 patients who have been systematically studied by Wolpe are very striking; he claims either 'cure' or 'much improvement' in 90 % of this group of cases. This work suggests that further examination of the uses of the term 'inhibition' in physiological psychology, and their systematic application to therapeutic problems, is a field likely to engage the attention of investigators of the future.

But it is not our concern here to go into all the details of the various uses of the concept in physiological psychology, though one other instance must be mentioned briefly. This is the notion of an inhibitory control by the higher centres over the lower centres of the nervous system—a formulation associated with the name of Hughlings Jackson. Head & Holmes (1911) were able to show how 'release of function', as Hughlings Jackson called it, may occur when there is a destruction of fibres leading from the cortex to the thalamus. There may ensue hyperaesthesias and other excessive responses to stimulation once this normally present inhibitory influence is withdrawn.* More recently Zangwill (1950, p. 137) has sought to explain the fear emotions in terms of 'diminished inhibition' in this Jacksonian sense. He argues that fear represents a scale which exhibits a progressively diminished influence of the higher centres and a 'corresponding release of the lower—and normally inhibited—neural activities'. In terms of this scale of intensity of fear we proceed from prudence and self restraint, through concentration and caution, apprehension and alarm, anxiety and anguish, to panic and finally terror. The earlier stages of this series are marked by a measure of cortical control and compensatory reactions in the face of increased interference with adjusted thought and action. The intermediate stages are characterized by the appearance of subcortically mediated activities, freed in part from cortical inhibition. Finally, there is a virtually uncontrolled discharge from the lower centres; the inhibitory influence of the cortex is largely out of action. Thus, when in a state of panic or terror, the human organism may react to a stimulus in a violent and uncontrolled manner somewhat after the fashion of the decorticate cat of Masserman's laboratory.

Inhibition, in the Jacksonian sense, implies some mechanism of control by the higher cortex. At this point it is appropriate to mention certain observations which have emerged from longitudinal studies of schizophrenic patients in which a markedly diminished inhibition of this kind seems to be operating. Thus Arieti (1955) in discussing the terminal stages of schizophrenia records a more or less uncontrollable tendency for the patients to place edible objects in the mouth. The contrary restraints which inhibit normal, civilized adults from doing this seem to be noticeably absent; the behaviour is strongly reminiscent of the food grabbing habits of infra-human

* Flugel has argued that the traditional distinction made by psychology between 'cognition' and 'orexis' is justified by neurological realities, and 'corresponds to some extent to a real difference of function between two major portions of the brain' (Flugel, 1933, p. 268).

organisms. Moreover, as the illness progresses the category of 'the edible' seems to disintegrate: the patients react to all objects of suitable size by an irresistible compulsion to place them in the mouth.* Rather similar behaviour can be produced experimentally, and Arieti discusses a series of experiments by Klüver & Bucy in the period 1937-39, when these investigators extirpated both temporal lobes of trained monkeys. Following this operation the animals exhibited a tendency to grasp, and place in the mouth, all objects of a suitable size within reach; all objects whether edible or not were treated in this way, and every operated monkey produced this reaction. Klüver (cited by Arieti) reports that the animals behaved as if 'acting under the influence of some compulsory or irresistible impulse... as if forced to react to objects'. Thus the Jacksonian conception of inhibition as involving restraint on lower brain processes whether by the temporal lobe or some other part of the cortex, is a notion which has possible applications to the normal psychology of emotion, to the understanding of the abnormal, and to comparative, developmental and physiological psychology. Rivers (1923) applies a rather similar concept to the study of dreams, and the early work of Silberer (1909) suggests it may be applicable to the understanding of the concrete modes of thinking which characterize the hypnagogic state. As Silberer when discussing hypnagogic imagery puts it: 'the tired consciousness, not having at its disposal the energy necessary to normal thinking, switches to an easier form of mental functioning', (Silberer, 1909, pp. 198-9). Similarly, Thurstone (1924), when discussing the difference between more intelligent and less intelligent forms of human thinking, appears to have in mind 'inhibition' in its Jacksonian sense when he argues: 'it requires inhibition of no mean order to retain a concept as such'. For Thurstone, as for Rivers and Silberer, inappropriate concretizations of thought are the products of relaxation of this inhibitory process. The possible applications of this concept of inhibition have been discussed in fuller detail elsewhere (McKellar, 1957).

Finally, there are a number of uses of inhibition in psychology which are perhaps best classed as miscellaneous; some emerge from the work of investigators of motivation who have been concerned with the interplay of impulse and inhibition. An illustration may be taken from Hertzberg's analysis of the psychology of philosophers. Hertzberg (1929) examined the biographies of a fairly large group of great philosophers in terms of their financial, occupational, marital and other adjustments: when compared with the norms for the population at large, he found them wanting. Hertzberg interprets this finding in terms of impulse and inhibition. He argues that it was not from weakness of affective impulse that the great philosophers of the past tended towards maladjustment; on the contrary, the evidence suggests that they were typically men of more than averagely strong impulse. The relative failure of their personal lives—and its by-product, the creation of systems of philosophy—resulted from the restraint of strong inhibitory processes upon these strong impulses. Thus, Hertzberg argues that, while ineffectiveness in action and adjustment *can* come from weakness of impulse, it can also come—and did in this case—when strong impulse is matched by equally strong inhibition. A rather different thesis is advanced

* We may note the resemblance of this to the behaviour of the 6-month to 2-year-old child. As Arieti puts it: 'the baby places in his mouth anything that comes within his reach, and licks, sucks or eats it'.

by Flugel (1934) in relation to similar issues. Flugel contends that strength of impulse and resulting pleasure can be *increased* by strength of inhibition; thus inhibition may be positively sought as a means towards greater satisfaction. It is in the context of the love relationship that Flugel provides his best-known analysis of this phenomenon. He has analysed the two conflicting conceptions of love—the Scandinavian and the French—as portrayed in Maurice Bedale's novel *Jerome*. In this the young Frenchman, Jerome, is wholly unable to understand the matter-of-fact conception of love held by his Scandinavian fiancée, bereft as this is of all the self-manufactured complications which, for the young Frenchman, add spice and interest to human love. Flugel's concept of 'greater satisfaction through inhibition', like Hertzberg's analysis of the psychology of philosophers, deserves fuller treatment as a study in certain complex aspects of human motivation than space allows. We mention these two studies as further illustrations of some of the interesting but miscellaneous applications of the concept of inhibition to psychological phenomena.

In conclusion, we would suggest for convenience the following names for at least some of the concepts of inhibition that have been employed in psychology:

- (1) Descriptive uses of the term, e.g. retroactive inhibition.
- (2) Pavlovian uses, e.g. 'internal inhibition', 'differential inhibition', etc.
- (3) Hullian uses, e.g. 'reactive inhibition', 'conditioned inhibition', etc.
- (4) Other uses of the concept in relation to the theory of learning, e.g. Harlow.
- (5) Eysenckian uses, which are derivatives of (2) and (3).
- (6) Psychoanalytic uses, e.g. in connexion with repression, the analysis of wit, and of the manic depressive psychoses.
- (7) Psychophysiological uses, as in the Jacksonian conception of inhibition.
- (8) Miscellaneous uses, as in Hertzberg and Flugel.

Clearly this classification is only one of a number of such possible classifications; moreover, we do not intend to imply that the uses discussed may not, on occasion, shade into one another. The aim of this preliminary analysis has been to emphasize the rich variety of uses the term 'inhibition' has had and still has in psychology. Yet all of these, in accordance with the *O.E.D.* definition quoted, seem to imply some form of preventing, hindering or checking of a psychological, physiological or psychophysiological process by a process of an antagonistic kind.

II. THE INVESTIGATION OF INHIBITORY PHENOMENA FROM THE STANDPOINT OF COMPARATIVE PSYCHOLOGY

By R. G. A. STRETCH*

Liverpool School of Tropical Medicine

It is appropriate for at least one paper of this Symposium to be devoted to a selective survey of recent studies in comparative psychology. Psychologists now know a great deal about the behaviour of animals under many different experimental conditions, and several important, theoretical concepts of inhibition have emerged from their extensive observations. In turn, these have proved to be of highest relevance to broad areas of human activity as, for example, the work of Eysenck (see Eysenck, 1957) and his colleagues testifies in relation to Hullian inhibitory theory. One aim of this paper is to discuss various features of infra-human behaviour where the concept of an inhibitory process has been employed, descriptively, to denote the characteristics of that behaviour. It is argued, for instance, that the phenomenon of habituation may be considered in this way; in addition, reference will be made to a number of physiological studies which relate to psychological concepts of inhibition. Secondly, much attention continues to be paid to investigation of the arousal and satiation of exploratory tendencies, spontaneous alternation, and 'perceptual curiosity'. Attempts have been made to account for these aspects of behaviour within the framework of the Hullian Two-Factor theory of Inhibition, and it is useful to inquire whether recent research continues to be incorporated satisfactorily, within this model.

We may begin with the problem of habituation. In 1906, Jennings (quoted by Sharpless & Jasper, 1956) drew attention to an interesting phenomenon of behaviour. If a shadow is passed over a sea anemone, the animal contracts rapidly. Repeated presentation of this stimulus at short intervals, however, leads to a temporary suppression of this reaction. It appears that some form of inhibitory control has developed which serves to check further response evocation. Harris (1943) finds that studies of habituation response decrement span the phylogenetic scale from unicellular organisms to the primates. Whilst he suggests: '... practically every stimulus-response relationship is subject to one or more types of decremental modification, in consequence of continued stimulatory activation' (p. 385), Sharpless & Jasper (1956) think it is unlikely that the same mechanisms mediate habituation in animals so distinct as amoeba on the one hand, and animals possessing a synaptic nervous system, on the other. In the latter case, it is possible to distinguish several forms of response decrement. These include effector fatigue, habituation as defined by Humphrey (1933), and Pavlov's internal inhibition. Habituation is distinct from effector fatigue in that a response which has become habituated to a specific stimulus may still be elicited by a novel stimulus, whereas a fatigued receptor will fail to respond to *any* stimulus of the same relative intensity. This phenomenon has yet to

* Formerly of the Psychological Laboratory, University of Sheffield.

be fully explained, however, at both the behavioural and physiological levels; perhaps its significance lies in the resemblance it bears to other inhibitory mechanisms, notably those described in terms of *stimulus satiation* (Glanzer, 1953), or *reactive inhibition* (I_R) (Hull, 1943). Indeed, Sharpless & Jasper (1956) suggest that the capacity of novel stimuli to capture attention (Berlyne, 1951; Broadbent, 1953), the phenomenon of curiosity (Berlyne, 1950, 1955), and insomnia produced by an unfamiliar environment, imply an initial sensitivity of the central nervous system to novel stimulation, and a loss of sensitivity as the stimulus becomes familiar. Stimulus satiation and reactive inhibition (I_R) may be viewed as a means of accounting for this loss of responsiveness; the former has been used, in recent years, to explain spontaneous alternation, and the latter, work decrement on monotonous tasks, and the decline of investigatory tendencies in a novel stimulus situation.

Although Sharpless & Jasper (1956) approached their study of the arousal reaction from the standpoint of neurophysiology, the results of their experiments have implications for this Symposium concerning: (i) the concept of disinhibition: (ii) the decline of exploratory tendencies; and (iii) the phenomenon of spontaneous alternation. The subjects in these experiments were thirty-five cats. Before the experiments were begun, electrodes were inserted at different levels of the brain stem: in the auditory cortex, subcortically in the region of the posterior hypothalamus, and in the mid-brain reticular formation. As a result of these procedures, it was possible to observe changes in the electrical activity of different parts of the brain, following afferent stimulation. In order to record and compare the patterns of activity produced by successive stimuli, it was necessary to establish a stable synchrony in the electrical activity of the brain; accordingly, the experiments were conducted after each animal had been induced to fall asleep. In the alert, waking state, low amplitude fast activity was observed, whereas in the sleeping animal, high amplitude slow waves were apparent. In a typical experiment, a sound of 500 cyc./sec. was presented to a sleeping animal for about 3 sec. This aroused the animal and produced a train of rapid, low-voltage activity on the electrocorticogram of approximately 3 min. duration. After the animal had again fallen asleep, the same stimulus was repeated with a similar result. After some thirty presentations, this reaction was markedly diminished, and the electroencephalogram tracing revealed arousal activity for no more than a few seconds. Presentation of a novel stimulus (a tone of 100 cyc./sec.) however, resulted in the reappearance of the original reaction.

In his classical studies of conditioning, Pavlov observed that the introduction of weak, unfamiliar stimuli often served to re-instate an extinguished response. This effect was termed *disinhibition*.* Sharpless & Jasper (p. 666) favour the view that disinhibition depends, not upon the 'blocking' of an inhibition, as this term implies, but upon the non-specific, or diffuse arousal effects induced by this stimulation. They support this interpretation by the fact that disinhibition was seldom observed in their own study if the effects of novel stimuli were allowed to dissipate before the familiar stimulus was repeated. In other words, after an animal had become

* In passing, we may note that the term 'dehabituation' has been used to refer to a similar effect in the case of unconditioned reflexes. This has been demonstrated, for example, after habituation of the withdrawal reflex in turtles (Humphrey, 1933), and after habituation of the startle reflex in rats (Prosser & Hunter, 1936). Quoted by Harris (1943) and Sharpless and Jasper (1956).

habituated to a particular tone, a different one was presented which led to a brief run of activation. When the cat had returned to sleep, the familiar stimulus was still incapable of eliciting arousal. It was observed, however, that a familiar stimulus might intensify the level of activity if it was applied whilst the animal was recovering from the effects of arousal by the novel stimulus.

It seems to me that there are similarities between this interpretation of disinhibition and the views expressed by Reid (1958) in this *Journal*. In this paper, the author reiterated a point made by Razran to the effect that the occurrence of a conditioned response may be strongly influenced by the animal's state of vigilance, and by its maintenance of posture, in the classical conditioning situation. It may be that any stimulus which revives vigilance will restore some of the familiar, contextual stimuli previously associated with the conditioned stimulus. When the conditioned stimulus is presented shortly afterwards, the probability of response evocation is consequently increased. In operant conditioning, as Reid's experiments with rats and pigeons have shown, disinhibition has proved difficult to demonstrate. Under these conditions, the animal is free to move about, and it is less likely that any stimulus will reinstate the distinctive context of the acquisition trials. In studies at Sheffield, it was equally difficult to observe disinhibition in an exploratory behaviour situation. Many experiments have shown that rats (and other mammals) will manifest extensive investigatory activity when they are placed in an unfamiliar environment. Using an enclosed T-maze, it was found, in prolonged trials of up to 90 min., that there are periods of several minutes duration when the rat is immobile (Stretch, 1959). By means of a lighting battery suspended over the maze-alleys, it is possible to induce a stimulus-change of brief duration by either switching on or switching off several small lights situated at equal distances along each arm. If this change is effected whilst the animal is immobile (usually towards the end of the trial), the rat appears to be momentarily aroused. In very few cases, however, does this evoke further ambulation. A single auditory stimulus of low intensity proved, likewise, to have little effect. It appears that merely alerting the animal is an insufficient condition for the re-arousal of exploratory behaviour. This may be contrasted with disinhibition reported in classical conditioning experiments, in which arousal might be expected to initiate trains of proprioceptive and interoceptive stimulation previously associated with the conditioned stimulus, thereby leading to a reappearance of the conditioned response.*

We may now turn to consider the application of inhibitory concepts to account for the decline of exploratory responses. Most mammals reveal, in their behaviour, a marked tendency to explore their surroundings; they tend to approach and enter every accessible part of their environment within a considerable area of their home-territory (Barnett, 1958). In doing so, they are often exposed to new objects and situations and many of their responses, as a consequence, become organized around novel features of their environments (Bindra, 1959). The study of exploratory

* I have been concerned with extinction rather than habituation and I am not at all sure of the relationship between the two. If habituation of exploratory behaviour is a process like operant extinction I should expect that sheer arousal would have little effect, but that the introduction of a stimulus that in its own right produces a component part of exploration, e.g. rearing up or sniffing, would be more likely to produce other parts of the exploratory pattern—(R. L. REID).

behaviour and related phenomena ('curiosity', Berlyne, 1950, 1955; manipulatory drives, Harlow, 1950; problem-seeking behaviour, Hebb & Mahut, 1955; Havelka, 1956; and visual exploration, Butler 1953) has received much attention in the last decade, and may be taken as evidence of a growing discontent with theories of motivation based upon the concept of biogenic or 'homeostatic' drives (Harlow, 1953*a, b*; White, 1959).

A central concept in this field, that of curiosity, was advanced by Berlyne in 1950. This model was expressed in postulate form as follows:

- (i) *When a novel stimulus affects an organism's receptors, there will occur a drive-stimulus producing response ($R \rightsquigarrow S_D$) (which we shall call 'curiosity') (1950, p. 73).*
- (ii) *As a curiosity-arousing stimulus continues to affect an organism's receptors, curiosity will diminish (1950, p. 74).*

To account for the satiation of curiosity responses, Berlyne employs Hull's Two-Factor theory of Inhibition. A temporary decline in curiosity is attributed to the accumulation of *reactive inhibition* (I_R), whilst a more slowly developing, permanent decline in responsiveness is interpreted in terms of the acquisition of a conditioned resting-habit or *conditioned inhibition* (sI_R). Berlyne demonstrated: (i) that rats enter an alcove in an exploratory box more frequently if it contains a novel object than if it is empty; (ii) that the amount of curiosity diminishes rapidly with continuous exposure to a novel object; and (iii) that there is a relatively permanent decrement in curiosity after initial exposure, that is, rats tend to enter the alcove less frequently as daily trials proceed. However, whilst Berlyne found evidence to support his theory, Montgomery (1953) did not. In his investigations at Yale University, Montgomery studied exploratory behaviour in simple mazes. Typically rats were placed in an enclosed 'H' or 'T' maze, the alleys of which were marked off in segments of 12 in. in length. The number of units an animal entered in each minute of a 10 min. trial, constituted a measure of exploratory activity. These animals entered many units initially, whereas few were entered towards the end of a trial. Montgomery (1953) also found that if a 10 min. trial was given each day, there was full 'spontaneous recovery' of maze exploration from one trial to the next. On the basis of this finding he concluded that conditioned inhibition was not necessary to account for the decline of exploratory tendencies because he could find no evidence of a permanent decrement in this behaviour. In reply to Montgomery's criticism, Berlyne stated: 'Conditioned inhibition might have subliminal value in certain situations, although the present state of Hull's theory unfortunately leaves us to ascertain which'. (1955, p. 238). This is most interesting, for I think it reveals that the concept of conditioned inhibition, at least in this context, is a relatively imprecise theoretical construct. Elsewhere in the same paper, Berlyne added: 'We can only hazard the suggestion that maze-runways present a large assortment of novel stimuli, so that a 10 min. exploration period may not allow enough time to satiate the curiosity aroused by most of them to the point at which a permanent decrement occurs' (1955, p. 245).

An experiment was undertaken, recently, to test this supposition (Stretch, 1959). In this, three groups of rats were allowed to explore a novel environment for different periods (20, 40, and 60 min. trials, respectively), on each of 4 successive days; they

were then returned to their home-cages for a period of 10 days before being given one further 20 min. trial in the maze. The results showed that the animals exposed to this environment for the longer periods during the first phase of the experiment, explored at a significantly lower rate during the test trial than the 20 min. group. Although these results support Berlyne's position, this conclusion was advanced tentatively because there seems to be some ambiguity in the statement of his theory. Berlyne suggests that the diminution of curiosity 'follows a course deducible from Hull's Two-Factor theory of Inhibition'. But Berlyne has offered an additional account of the decline of curiosity responses. He wrote: '...we cannot dismiss this decline glibly as nothing but extinction of exploratory responses. A response should extinguish only if it ceases to be reinforced, and without satiation of the curiosity drive it is difficult to see why this should occur. If however, we assume that the curiosity drive aroused by a particular object diminishes by being satiated, then the decline in exploration is predictable. A decrease in drive level will first of all mean a decrease in the reaction potentials of exploratory responses (Hull, 1943, 1952), *and then finally, when there is no drive left to be reduced, the responses will extinguish because they are no longer followed by drive-reduction.*' (1955, p. 238; italics mine.)

If there is ambiguity in this passage, it may be said to arise as a consequence of the phrases which have been placed in italics. In Hullian theory, reaction potential is alleged to equal drive multiplied by habit strength, according to the equation:

$${}_sE_R = D \times {}_sH_R \quad (\text{Hull, 1952}).$$

Thus, when a drive is reduced to zero, as in Berlyne's statement, reaction potential also equals zero. Under these circumstances, it is difficult to envisage how extinction might occur since responses can no longer be elicited, and therefore, inhibitory aggregate ($\dot{I}_R$) cannot be held to increase. On these grounds, it seems that the application of Hull's inhibitory concepts (I_R and ${}_sI_R$) to account for the decline of exploratory responses is not necessary in spite of experimental evidence to the contrary (Berlyne, 1955; Glickman, 1958; Stretch, 1959). We may also note that Berlyne postulates (1954) have criticized this type of formulation on the grounds that, if novel stimuli elicit this drive, it would mean that the drive is produced *after* the response has been made which it is supposed to be motivating. Brown (1953) has raised a similar objection to Harlow's (1953a) concept of 'exteroceptively aroused drives', i.e. manipulation and visual exploration (Butler, 1953). In Butler's experiment, monkeys learned to manipulate a latch in order to see out of an opaque box in which they were housed. It is unlikely that the drive motivating this manipulation is dependent upon an increase in stimulation; if visual exploration was the motive, 'then the monkeys must have been unmotivated until the window was opened after a correct response' (Brown, p. 54).

A different view of exploratory behaviour has emerged more recently, in the writings of Hebb & Thompson (1954), Hebb (1958), Leuba (1955), and Glickman (1958) which seems to overcome this difficulty. It suggests that animals will seek conditions that produce a moderate level of arousal and will tend to avoid those that lead to monotony (low arousal) or pain (high arousal). This view implies that we may regard exploratory behaviour as an attempt on the part of an animal to maintain

an optimal level of excitement. It would seem to follow that depriving an animal of one of his main sources of sensory input would lead him to explore a novel environment more than a control animal in order to attain an equivalent degree of arousal. Glickman (1958) restricted the peripheral vision of a number of hooded rats, and found that they explored a simple maze significantly more than his control group. He concluded this result supports a concept of optimal arousal. Furthermore, he found that the exploratory activity of the experimental and control animals declined in a way that was consistent with Berlyne's hypothesis. This is of particular interest because Glickman suggests that a neurophysiological basis may be found for Berlyne's inhibitory concepts in the experiments by Sharpless (1954) and Sharpless & Jasper (1956). To recapitulate a little, these writers demonstrated that, with repeated presentation, an auditory stimulus loses its capacity to arouse the ascending reticular formation of the brain-stem; this occurred in a way which relates closely to Berlyne's inhibitory concepts. In discussing the duration of states of habituation, Sharpless & Jasper distinguish between 'phasic' arousal which barely outlasts the stimulus, and is highly resistant to habituation, and 'tonic' activation which may persist for several minutes but is readily susceptible to habituation control. They suggest that tonic arousal is probably mediated by the lower portions of the ARAS*, whilst the phasic arousal response may be mediated by the upper regions of the system (probably the diffuse thalamic projection system). They go on to say: 'We were struck by the fact that the animals seemed less responsive generally and became habituated more rapidly to experimental stimuli on the second and third days of recording than on the first. This seemed to indicate... that a state of habituation to the experimental situation developed which persisted from one day to the next. The effect was so marked in some cases, that it was necessary to interpose several day's rest between successive recording sessions. Given such a rest, the animal's normal responsiveness was usually restored' (1956, p. 664). The similarity between this phenomenon and the more permanent decremental effects which have been observed in experiments on exploratory behaviour seems obvious enough. Glickman (1958) suggests that the course of habituation in these experiments is 'functionally similar to the apparent decline in "excitation power" of the open field or Y-maze situations' which have been used to measure exploratory activity, and according to Beach (1957), Sharpless (1954) proposed that excitation of the reticular substance produces a 'dynamogenic' or facilitatory effect on new processes generally, and that this was the basis for exploratory behaviour. It seems clear we may suggest that 'response to novelty' and the subsequent decline in responsiveness depends upon the integrity of the ascending reticular formation.

Furthermore, Stretch & Kilpatrick (1959) have recently investigated the effects of a central depressant (sodium amytal) and a central stimulant (amphetamine) upon curiosity and exploratory behaviour. These drugs were chosen because they may act directly upon this part of the central nervous system (Bradley, 1957); in addition, their use enabled us to examine Eysenck's (1957) drug postulate. In this study, we were concerned with the concept of an optimal level of arousal or 'excitement'. If a depressant drug raises the threshold for arousal, we might expect an

* Ascending Reticular Activating System.

animal to engage in *more* investigatory activity than a control animal in order to attain an equivalent degree of arousal; conversely, a reduction in this threshold through the action of a stimulant drug may mean that an equivalent degree of excitement can be derived from *less* exploration than the control group. This hypothesis was indirectly confirmed in two situations: a maze-environment, and an exploratory box in which the time spent sniffing at a novel stimulus card was the dependent variable. In one case, the amytal group manifest significantly more, and the amphetamine group significantly less, exploration and curiosity than the control group. Secondly, as Eysenck's (1957) drug postulate suggests that depressant drugs increase cortical inhibition, and that stimulant drugs decrease it, we predicted that the *rate of adaptation* to novel stimulation (as a measure of response-induced inhibition (I_R)) would be most rapid for the amytal group and least rapid for the amphetamine group. This prediction was also confirmed at a satisfactory level of significance, and we may note the parallel between this experiment and an investigation by Sinha, Franks & Broadhurst (1958) in which the effects of a stimulant and a depressant drug upon a measure of spontaneous alternation were shown to support Eysenck's proposition. Further work is being undertaken to examine these relationships using other drugs and different dosages.

Not long ago, Estes (1956) remarked that: 'Those most active members of the Hullian family, the brothers inhibition, seem to be drifting into different social strata. Reactive inhibition (I_R) has come into very wide usage as a label for any sort of temporary response decrement; . . . Conditioned inhibition (sI_R) remains an ostentatiously theoretical construct.' This view is particularly apt with regard to the application of these concepts to explain the diminution of curiosity responses judging by Montgomery's (1953) criticism of Berlyne's theory, and the latter's reply. In studies of alternation behaviour, reactive inhibition has been employed, by itself, to account for this phenomenon. In these experiments, an animal is permitted to make at least two successive choices between the arms of a 'T'-maze; in doing so, it usually displays a marked tendency to choose a different arm on the second trial from its first (Glanzer, 1958). Whilst the rat has been the most popular subject for these studies, Lepley & Rice (1952) have observed a similar tendency in the behaviour of *paramecia*. Grosslight & Ticknor (1953) have also reported this phenomenon using mealworms (larvae of *Tenebrio molitor*) as subjects. These authors rather uncritically accepted the concept of reactive inhibition as a means of explaining alternation behaviour, and we may note that Rice & Lawless (1957) failed to confirm this hypothesis with *Planaria dorotocephala*. However, this mode of explanation has been commonly used to account for alternation in the rat (see Barnett, 1958). It states that the most recent response is associated with the greatest primary inhibition (I_R), and therefore, is not likely to be repeated in competition with a less recent response—at least, not until reactive inhibition from the first response has dissipated. Thus, Heathers (1940) showed that alternation decreased with an increase in running time; if the interval between the two choices was a matter of 15 sec., alternation occurred on approximately 80 % of the trials, but if the interval increased to 120 sec., the percentage dropped to the region of 65 %. Several studies have challenged this notion. Using a cross-shaped maze, Glanzer (1953) and Montgomery (1952) both found it necessary to reject this explanation because they discovered that a rat *will* repeat a recent

response in order to encounter less familiar stimuli. Glanzer postulated a new concept, namely, *stimulus satiation* to explain his results, and Montgomery interpreted the effect in terms of his *exploratory drive*. Moreover, Estes & Schoeffler (1955) have shown that alternation falls to a chance level when, after forced trials to one arm of a 'T'-maze, the maze is rotated through ninety degrees; in other words, extra-maze stimuli appear to be important determinants of this behaviour. The criticism that has been levelled against reactive inhibition as an explanation for alternation behaviour may not be thoroughly legitimate through its lack of sophistication with regard to the definition of response. Some investigators (Glanzer, 1953; Montgomery, 1952*a*) have presented data which is inconsistent with *one* interpretation of the Hullian theory, namely, that primary inhibition (I_R) accrues as a function of the overt response, that is, the act of turning right or left at the choice point. Their experiments do not necessarily present a crucial difficulty for the Hullian model; as Broadbent (1959) has suggested, one *might* argue that: 'it is the visual response, internal and unobservable, which is subject to such inhibition. Such a re-interpretation is in line with the approach of Berlyne (1951), and seems to be perfectly valid.' We may note, however, that if reactive inhibition is treated in this way, the grounds for differentiating between this concept and stimulus satiation are reduced.

Of greater significance, is the study by Walker, Dember, Earl & Karoly (1955*a*), which showed that the effect of alternation was to vary sensory input in successive choices; this applied to both internal and extra-maze stimuli. In a second study, Walker *et al.* (1955*b-c*) showed that alternation may partly depend upon the nature of the response. By complicating the entries into each goal-arm of the 'T', they demonstrated that response alternation does occur. This experiment suggests, however, that 'an important factor is the readiness with which two alternatives can be discriminated; merely turning left or right are not very different, but one set of conditions is readily distinguished from another' (Barnett, 1958). Walker (1958) has recently postulated a new concept, *action decrement* (I_c), to account for spontaneous alternation. Action decrement is used to refer to a state similar to Sharpless & Jasper's (1956) *habituation*. It is a form of inhibition which checks or delays the performance of a recently practised, new response, in order to protect the consolidation process of learning from interference. Provision is made in this theory, by reference to the activity of the reticular formation, for motivational variables, so that predictions can be derived concerning the effect of reward upon alternation tendencies (Walker, 1956). Unfortunately, shortage of space prevents a more detailed exposition of this model.

In conclusion, it is unwise to ignore the fact that there are many inhibitory phenomena which are not necessarily incorporated within a specific theory of behaviour. For example, a state analogous to Pavlov's internal inhibition has been 'artificially induced' in some sense, by Bures, Buresova & Zahorova (1958).

These investigators, working in Prague, have shown that a *spreading of cortical depression* can be induced by the application of potassium chloride to the surface of the cortex. By means of a small trephined hole in a rat's skull, it is possible to place a tiny piece of filter-paper, soaked in potassium chloride solution, directly on to the surface of the cortex. In their experiments, these workers established certain conditioned responses in rats. After a 25% solution of potassium chloride was

applied over an area of 2 mm. in diameter, the rats were replaced in the conditioning apparatus, and the previously established conditioned responses were re-tested. The results showed that 1 min. after the application of the solution, over either the visual or motor cortex, conditioned responses could still be elicited. However, several minutes later, and for up to 3 hr., conditioned stimuli failed to evoke the appropriate response. Electrical changes of comparable duration occurred in the cerebral cortex under these conditions; a deep depression of E.E.G. activity was manifest and lasted for several hours. This particular study has been mentioned, because this technique may afford a most useful research tool for the investigation of classical conditioning phenomena.

Some reference has been made to the ascending reticular activating system in this paper because much recent physiological research suggests that it plays an important role in the arousal and satiation of certain responses. An attempt has been made to show the relevance of this work to psychological concepts of inhibition, with particular reference to the study of exploratory behaviour.

III. EFFECTS OF DRIVE AND EFFORT ON INHIBITION WITH REINFORCEMENT

By DONALD C. KENDRICK*

Institute of Psychiatry, Maudsley Hospital, University of London

I. INTRODUCTION

The literature concerning the phenomenon of inhibition with reinforcement, having been discussed elsewhere (Kendrick, 1958a), will not be reviewed again. It has previously been shown that if rats are trained to run a runway for a water reward there will come a time when they will refuse to run even though they are very thirsty. This fact has been explained as a function of massed practice producing reactive inhibition (I_R), the dissipation of which results in the conditioning of a negative habit, conditioned inhibition (sI_R), which has properties of a positive habit, e.g. it can be experimentally extinguished. Hull (1943) was very inconsistent in his formulation of the concept of sI_R , as has been pointed out by many writers (e.g. Osgood, 1953; Iwahara, 1957; Jones, 1958) and the present writer has used Jones's reformulation of the Hullian learning equation, for which he has found some experimental support (Kendrick, 1958a). In considering the inconsistencies, the writer refused to accept the prediction made by Gleitman, Nachmias & Neisser (1954) from Hullian learning theory, from which all this work began. The prediction was rejected because it contained Hull's inconsistencies; moreover, the experimental findings did not support the mechanism of extinction that those authors envisaged. Instead the writer used many modifications of the Hullian learning system to explain his experimental findings. The production of I_R was considered a central phenomenon; a revised mechanism of sI_R production was utilized (Eysenck, 1956); the concept of tolerance levels for I_R was also adopted (Kimble, 1949; Ellis, 1953), the tolerance being dependent upon the strength of the primary drive in operation at that particular time; and the reformulation of the Hullian learning equation was also adopted (Jones, 1958). Despite these revisions certain predictions remained unvalidated. From the tolerance hypothesis it is predicted that, all other things remaining equal, under conditions of low drive the tolerance level for I_R will be less than for that under conditions of high drive. Thus in the phenomenon being investigated animals trained to run a runway under low drive should extinguish faster than a comparable group run under conditions of high drive. Further, from the reformulation of the learning equation it follows that if the animals are run to the point where the strength of sI_R counter-balances the strength of sH_R or in fact outweighs it, then any increase in drive will only serve to increase the potentialities proportionally, as after a rest period when all I_R has dissipated both the positive and negative habits will be activated by the same primary drive. Therefore there will be no recovery of the habit with increased drive, a prediction which could not be made from Hull's original learning equation.

* This study was made possible by a Bethlem Royal and Maudsley Hospitals' Research Grant.

It was postulated (Kendrick, 1958*b*) that as Landfried & Wike (1956) could not produce inhibition with reinforcement when a runway 46 in. long was used that it may have been that the difference between 10 ft. and 46 in. may be critical in the production of sI_R . Fewer responses are required to run the shorter distance; therefore the central nervous system received less stimulation and so less I_R developed. It was necessary to make a controlled study to test this hypothesis. Three predictions were made: (1) Under low drive conditions extinction would be significantly more rapid than that produced by high drive conditions. (2) That performances under low drive conditions would be significantly worse than under high drive conditions because low drive would produce a lower sE_R . (3) Conditions of low effort would significantly defer the point of extinction compared with high effort conditions.

II. EXPERIMENT

(a) *Experimental design*

Three groups of three male rats each were formed.* Group A served as a control group for group B, and were trained to run the runway, 10 ft. long, under a water deprivation level of 23 hr. Group B served as a control group for group D and were trained in exactly the same manner as group A except that they had a water deprivation level of 1 hr. Group D was tested under the same conditions as group B except that the runway length was only 4 ft. long. Thus two levels of deprivation were used, and two levels of effort, effort being experimentally defined in terms of runway length.

It was proposed that all subjects should run 25 trials a day, the trials being massed. On each trial the animal was rewarded with a 0.25 c.c. of water. To receive the reward the subject had to run the length of the runway then enter a goal box, depress a level bar three times with a downward thrust of 60 g., and then drink the reward from a reinforcement dish. Each subject was allowed 20 c.c. of water per day, 6.25 c.c. being consumed during the testing period, and the rest was used to manipulate the drive level. Thus group A received 13.75 c.c. of water 23 hr. before being tested; if a subject refused a trial then it was compensated at the rate of 0.25 c.c. per trial with the 13.75 c.c. of water. Groups B and D received 13.75 c.c. of water 1 hr. before being tested; if any animal in these groups refused trials they were also compensated in a similar manner, but the compensation was given 1 hr. after testing and not with the pre-testing water. This manipulation of the water was done so that all animals should be equated for h_t (Eisman, 1956), but should differ significantly on drive level at testing as the rat's stomach is empty in approximately 4 hr. As many relevant variables as possible were controlled for this experiment. Pre-testing experience was standardized, also animal husbandry, the light/dark cycle, colony and test room temperatures, days to weaning and to caging singly, cages, diet, and 'gentling'. Two weeks before testing, the animals were gradually brought onto a 23 hr. water deprivation schedule, the deprivation level being increased daily by 1 hr., from 16 to 23 hr. All animals were trained on 23 hr. water deprivation and all used the full runway length during training; this was done to equate for learning.

There were two criteria to be attained before an animal was considered to have extinguished. First, it must refuse to press the bar for the third time within 3 min. of the start box-door being opened, twice within the first five trials, and secondly it must do this on two successive days. If an animal refused for two consecutive trials

* Actually four groups were tested, but the results of one group are irrelevant to this paper.

after the first five trials had been completed then testing was discontinued for that day, but was not counted as extinction.

Thirty-six days after extinction the animals were tested for spontaneous recovery, for recovery of the habit under increased drive, i.e. groups B and D were tested under 23 hr. water deprivation to test the reformulation of the learning equation.

(b) Apparatus

A horseshoe-shaped runway was constructed in six separate pieces. The internal measurements of the runway were 4 in. wide and 6 in. high. The floor, ceiling and outside walls were constructed of 16 G black anodized aluminium. The inside wall was made out of opaque Perspex, so that as the runway was internally illuminated, the shadow of the subject could be seen all the time it was in the runway. The sides were positioned in grooves on the floor and ceiling so that they could be moved back, and forwards to allow a subject to be picked out of the runway easily if it refused to move. As the runway was in six sections all joins were overlapping. The six sections comprised two 1 ft. straight sections which served as starting and goal boxes, two 3 ft. straight sections and two curved sections with an inside radius of 1 ft 3 in. giving a nearside distance of 2 ft. The runway's ceiling had a pair of 2.5 bulbs placed at 1 ft. intervals, so that there were a pair in the starting box placed centrally and similarly in the goal box, three pairs in the 3 ft. sections and two pairs in the 2 ft. sections. The starting box had a door which was pivoted vertically and which opened out into the runway on the off side. The door was operated manually and the handle of the door, which protruded through the top of the start box, slotted into a clip and depressed a micro-switch, starting a stop-watch which timed the animal until it made its first press on the bar, at which point another watch was set in operation which recorded the time the animal took to make the three presses. The goal box had a similar door to the start box but had no timing apparatus connected to it, and was only used during training and post-extinctively. The goal box was made into a Skinner type device. A lever 2 in. long, protruding 1 in., was placed centrally on the end wall of the goal box 1 in. from the floor. A guarded nozzle was placed in the adjacent wall 1 in. from the ground beneath which was placed a watch-glass 1½ in. in diameter directly beneath it, into which 0.25 c.c. of water was expelled from the nozzle. The lever pressing mechanism was so designed that one, two or three presses could be made before the water dropped into the reinforcement dish. Above the nozzle a 2.5 bulb was placed which only lit up when the correct number of presses had been made. When the correct number of responses had been made the circuit went dead so that the subject could not be reinforced more than once on any trial.

(c) Subjects

The subjects consisted of nine male albino rats, taken from an original group of twelve, the three remaining being tested, under conditions irrelevant to this paper. The subjects belonged to a pure strain. Each subject had the same father, and the mothers were interuterine sisters and the interuterine sisters of the father. When testing began the age range of the animals was from 78 to 82 days. On the day before testing began it was found that the mean weight of each group (each animal having been randomly assigned to a group), did not differ significantly from the mean weights of the other groups, nor did the mean intake of water after 23 hr. water deprivation differ between the groups. No illness was discovered during any part of the experiment in any of the animals, and a mean body weight of 18 g. was recorded during the testing period.

(d) Procedure

Training was divided into three sections: habituation, training and testing. The animals were not assigned to groups until the first day of testing. Before habituation began the animals had gradually been brought onto the 23 hr. water deprivation level over a period of 2 weeks, the deprivation level having increased by 1 hr. each day

from 16 to 23 hr. The animals were also gentled each day over this same period for 5 min.

Habituation. For two days each subject was allowed to explore the runway freely for 15 min. The apparatus was set in working order, with the bar-pressing mechanism set so that one press would produce a reward. Both gateways into the runway were open.

Training. A series of criteria had to be attained by each subject in this period. All subjects had to reach the same criterion before they moved on to the next stage, so that the first animal to reach a criterion had to wait for the others to catch up; this was to control for the ageing factor. On the first day the subject was placed in the goal box with the door into the runway closed. The bar was set so that one press would produce a reward. The bar itself was weighted with 50 g. so that a pressure of 10 g. would operate the lever. Criterion 1: the subjects had to learn to press the weighted bar 10 times within 7 min. and drink after each press. Criterion 2: the bar had to be pressed 10 times in 7 min. with the weight removed. Criterion 3: the subject had to learn to run from the starting box to the goal box and press the lever once 12 consecutive times in 20 min. Criterion 4: the animal had to run the runway and press the bar twice for reward for 5 trials and then press it three times for 20 trials in unlimited time.

Testing. Testing day 0 was given which consisted of all the animals being run for 25 trials under 23 hr. water deprivation. The animals were then randomly assigned to one of three groups, each group consisting of three subjects. On testing day 0 the mean pressing and running times per group were found not to differ significantly from one another; it was therefore assumed that all subjects began testing at approximately the same level of proficiency. On testing day 1 the subjects were tested under their various group regimes of drive and effort. Testing was simple. A subject was placed in the starting box and the door immediately opened; the subject ran to the goal box, pressed the bar, and as soon as it had finished drinking it was lifted out and returned to the starting box and another trial immediately started. This was repeated for 25 trials. No animal was rewarded if it refused to run or press the bar.

III. RESULTS

The criteria of extinction were met by all subjects within a range of 13–27 days. Analysing days to extinction first it was found that group B did not extinguish significantly faster than group A (group A, $\bar{X} = 25.00$; group B, $\bar{X} = 17.67$; $t = 2.16$, with 4 d.f. However, it can be seen that the results are in the predicted direction, and the insignificance was probably due to a greater variance in group B. Group B did not differ significantly from group D either (group B, $\bar{X} = 17.67$; group D, $\bar{X} = 15.00$; $t = 0.79$, with 4 d.f.). Group D did however extinguish significantly faster than group A ($t = 6.47$, sig. at 0.01 level, with 4 d.f.). This result is the more surprising as the prediction was that group D would extinguish after group B. Considering trials to extinction the results are more conclusive. Group B required significantly less trials to meet the criteria of extinction than group A, (group A, $\bar{X} = 576.67$, group B, $\bar{X} = 359.67$; $t = 5.11$ sig. at 0.01 level, with 4 d.f.). There is no difference between groups B and D (group B, $\bar{X} = 359.67$; group D, $\bar{X} = 317.33$; $t = 0.97$, with 4 d.f.). It is concluded therefore that the prediction that extinction will be brought about more rapidly under conditions of low drive has satisfactorily been supported, but no evidence has been found to support the prediction that low effort will decelerate the extinction process.

The median running and pressing times were calculated for each subject per day, and as in the previous experiment (Kendrick, 1958*a*) the data have been divided into two sections, in order to equate each subject's learning curves. The first section, period 1, the learning period, contains the data for each subject from the first six testing days, and the second section, period 2, the extinction period, contains the

data from each animal, counting back from each individual extinction day to the 6 preceding days; thus 12 days' testing is used in all, and as the range of days missed out is from 1 to 15, the median running time scores from these days have been added together per group and a mean median running time calculated in the period which may be termed the asymptotic period.

Table 1. *Summary of analysis of variance.*
Median running times. Group A $\times$ group B

Source	D.F.	M.S.	F
Periods (<i>P</i>)	1	0.01	0.5
Days within <i>P</i> (<i>D</i>)	10	0.05	2.5*
Groups (<i>G</i>)	1	0.23	11.5**
Subjects within <i>G</i> (<i>S</i>)	4	0.03	1.5
<i>P</i> $\times$ <i>G</i>	1	0.00	0.00
<i>S</i> $\times$ <i>P</i>	4	0.01	0.5
<i>S</i> $\times$ <i>D</i>	4	0.01	0.5
<i>D</i> $\times$ <i>G</i>	10	0.01	0.5
Residual	40	0.02	—
Total	71	—	—

*Significant at 0.05 level. ** Significant at 0.01 level.

Table 2. *Median running times. Group B $\times$ group D*

Source	D.F.	M.S.	F
<i>P</i>	1	0.00	0.00
<i>D</i>	10	0.05	5.00**
<i>G</i>	1	0.00	0.00
<i>S</i>	4	0.31	31.00**
<i>P</i> $\times$ <i>G</i>	1	0.00	0.00
<i>S</i> $\times$ <i>P</i>	4	0.025	2.5
<i>D</i> $\times$ <i>G</i>	10	0.01	1.00
Residual	40	0.01	—
Total	71	—	—

** Significant at 0.01 level.

Table 3. *Median running times. Group A $\times$ group B*

Source	D.F.	M.S.	F
<i>P</i>	1	0.52	8.76**
<i>D</i>	10	0.048	0.06
<i>G</i>	1	0.00	0.00
<i>S</i>	4	0.06	1.00
<i>P</i> $\times$ <i>G</i>	1	0.04	0.67
<i>S</i> $\times$ <i>P</i>	4	0.02	0.33
<i>D</i> $\times$ <i>G</i>	10	0.007	0.11
Residual	40	0.06	—
Total	71	—	—

** Significant at 0.01 level.

Taking each subject's median running and pressing times in periods 1 and 2 the groups have been compared by analysis of variance techniques; see Tables 1-4 for the design and results of these analyses.

Table 4. *Median pressing times. Group B $\times$ group D*

Source	D.F.	M.S.	F
<i>P</i>	1	0.42	26.25**
<i>D</i>	10	0.026	1.63
<i>G</i>	1	0.125	8.13**
<i>S</i>	4	0.125	8.13**
<i>P</i> $\times$ <i>G</i>	1	0.01	0.63
<i>S</i> $\times$ <i>P</i>	4	0.02	1.25
<i>D</i> $\times$ <i>G</i>	10	0.008	0.5
Residual	40	0.016	—
Total	71	—	—

** Significant at 0.01 level.

Considering running times first, two group comparisons have been made: A with B, and B with D. The comparison of group A with group B shows a very clear-cut result; group B ran consistently slower than group A, thus supporting the second prediction that low drive would result in lower performance. The comparison between groups B and D is singularly disappointing, and the results are mainly negative; there is certainly no support for the third prediction. The bar pressing time results of the comparison of groups A and B show no group differences in performance. This is not surprising, however, as this habit would almost certainly be stronger than the running response, because the reward is more immediate to the bar response, and the response itself is much less conducive to individual variation. There is a period difference which will be found to be consistent throughout all the pressing time results, and that is that in the period 2 the times are faster than in period 1. Groups D and B show this same pattern. The group difference between B and D cannot be taken seriously as the subject within group mean square variance is so large, and the group mean square variance when tested against this variance is 0.054 which is clearly not significant.

The data were also examined for the evidence of warm-up effects at the beginning of each day's set of trials and for the slowing down of time at the end of a set of trials, the later phenomenon, it is postulated, being due to the depressing action of I_R . To observe these phenomena the first five and last five trials a day have been used. The first five trials per subject per day have been added together and a general mean calculated. The last five trials were treated similarly. Both measurements have been analysed in the same way. See Table 5 for the design and summary of the results. In the analyses where the interactions are not significant they have all been summated with the second-order interactions and used as one residual to test the main effects. The analyses of running times revealed no differences between group A and B, nor between groups B and D. The bar-pressing speeds showed a period difference in the comparison of groups A and B, the times showing the characteristic speed up in period 2; a similar result was found in comparison between groups B and D.

One further analysis was carried out to see if the running times in the asymptotic period were different from those in period 2; the results were negative.

Post-extinctively it was found that there was no spontaneous recovery of the response, that both habits were extinguished, and that an increase in drive did not

Table 5. *Summary of the analyses of variance of the means of the first and last five trials per day, first block group A $\times$ group B, second block group B $\times$ group D*

Source	D.F.	M.S.V.	F	M.S.V.	F
Running times					
Groups (<i>G</i>)	1	0.46	1.80	0.00	0.00
Sections (<i>SC</i>)	1	0.04	0.16	0.22	2.42
Periods (<i>P</i>)	1	0.01	0.04	0.00	0.00
Days within <i>P</i> (<i>D</i>)	10	0.052	0.21	0.059	0.65
<i>G</i> $\times$ <i>SC</i>	1	0.02	0.71	0.02	1.05
<i>G</i> $\times$ <i>P</i>	1	0.03	1.07	0.02	1.05
<i>G</i> $\times$ <i>D</i>	10	0.03	1.07	0.02	1.05
<i>SC</i> $\times$ <i>P</i>	1	0.14	1.21	0.16	8.42*
<i>SC</i> $\times$ <i>D</i>	10	0.005	0.18	0.011	0.53
<i>G</i> $\times$ <i>SC</i> $\times$ <i>P</i>	1	0.00	0.00	0.00	0.00
<i>G</i> $\times$ <i>SC</i> $\times$ <i>D</i>	10	0.028	—	0.019	—
<i>G</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—
<i>SC</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—
<i>G</i> $\times$ <i>SC</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—

* Significant at 0.05 level.

Pressing times					
<i>G</i>	1	0.00	0.00	0.09	1.96
<i>SC</i>	1	0.04	0.52	0.06	1.30
<i>P</i>	1	0.47	6.10*	0.45	9.78**
<i>D</i>	10	0.053	0.69	0.026	0.57
<i>G</i> $\times$ <i>SC</i>	1	0.00	0.00	0.00	0.00
<i>G</i> $\times$ <i>P</i>	1	0.02	1.67	0.01	1.67
<i>G</i> $\times$ <i>D</i>	10	0.02	1.67	0.01	1.67
<i>SC</i> $\times$ <i>P</i>	1	0.02	1.67	0.01	1.67
<i>SC</i> $\times$ <i>D</i>	10	0.01	0.83	0.01	1.67
<i>G</i> $\times$ <i>SC</i> $\times$ <i>P</i>	1	0.00	0.00	0.01	1.67
<i>G</i> $\times$ <i>SC</i> $\times$ <i>D</i>	10	0.012	—	0.006	—
<i>G</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—
<i>SC</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—
<i>G</i> $\times$ <i>SC</i> $\times$ <i>P</i> $\times$ <i>D</i>	0	—	—	—	—

* Significant at 0.05 level. ** Significant at 0.01 level.

disinhibit those groups extinguished under conditions of low drive, thus supporting the reformulation of the Hullian learning equation.

IV. DISCUSSION

It might appear that the effect of low drive on performance and toleration level might be confounded, i.e. the faster extinction under low drive may be due only to the fact that performance is not so good. This however is not the case. Hull's fourth postulate (1943) states that habit strength is dependent upon the number of reinforcements and not drive level; it is the reaction potentiality which is dependent upon drive level. As groups A and B were shown not to differ significantly on testing day 0 and as drive at this point was constant, it must be assumed that at this point habit strength is the same in both groups, and will remain roughly equivalent throughout the testing as reinforcement is kept constant. The mechanism described as producing this extinction is that where habit strength is equalled or outweighed by conditioned

inhibition; the strength therefore of reaction potentiality does not enter into the question. If then group B extinguishes in less trials and probably less time than group A some other factor must be increasing the effectiveness of the conditioned inhibition, and a theory of reduced tolerance with low drive suitably accounts for the experimental facts, and it is therefore considered that the toleration hypothesis has been supported.

It is necessary, of course, to repeat the experiment concerning the role of effort in the production of reactive inhibition, as this function is notoriously volatile. The results of Aiken (1957) have shown the complexity of this problem. The present experimental results are in line with the assumption made by Ellis (1953) that the critical level for reactive inhibition is independent of the effort variable. This hypothesis is further borne out by the fact that there are no group/section differences between groups B and D on running times. If effort had had an effect it would be expected that group B would show greater warm-up effects in section 1, period 2, than group D.

The fact that there is no recovery of the response by groups B and D when they were tested under 23 hr. water deprivation is a result that could not have been predicted from the original Hullian learning equation, but which lends validity to the reformulated equation already discussed. The lack of spontaneous recovery on the part of all animals shows that this extinction cannot be regarded as a normal experimental extinction phenomenon.

The writer has also shown (Kendrick, 1959) that a group of animals trained to run this alleyway under identical conditions to group A but only being reinforced randomly 15 out of every 25 trials extinguish significantly faster than group A. Moreover, there was no indication that a non-reinforced trial produced any form of disinhibition. It has therefore been concluded that Hovland's (1936) concept of inhibition of reinforcement cannot be supported. The fact that this group extinguished significantly faster than group A would seem to be the product of two factors; (a) a lower strength of sH_R and therefore correspondingly less sI_R necessary to equal it; and (b) non-reinforcement has either increased the production of I_R or a type of inhibition as Hull suggested, *extinctive exhibition*, which interacts with I_R .

Eysenck (1956) postulated that sI_R was maximally accrued when involuntary rest pauses occurred. This phenomenon was produced by the tolerance level for reactive inhibition having been reached, this causing an automatic resting response. It is still postulated by others (Starr, 1958; Kendrick, 1959) that enforced rest pauses produce a similar effect, though the actual amount of sI_R accrued may not be so great under these conditions. It would be very difficult under the conditions of this experiment (pursuit roter) to pick out what might be considered as an I.R.P. The writer concludes that I.R.P.'s are not necessary for the production of sI_R , but this does not imply that I.R.P.'s may not generate a greater increment of sI_R .

There appears to be no good reason why sI_R should not equal sH_R . It is pointed out by Reid in this symposium that the asymptote of sI_R may not be as high as the asymptote of sH_R . The asymptote of sI_R from the evidence of the number of successive refusals by the subjects, appears to equal or in fact be stronger than sH_R as at the beginning of a day's testing both habits are activated by the same drive (B).

The effectiveness of sI_R will depend upon the tolerance level reached; if this is low then sI_R will be effective.

From the results of this work it is concluded that the hypotheses for inhibition with reinforcement being either stimulus or reinforcement produced have not been substantiated, and that the concept of antagonistic habits sH_R against sI_R deals more adequately with the results.

I must now turn to answer some of Reid's criticisms of the theoretical model that I use to explain my results. It is obvious that Reid thinks that I use the Hullian learning theory as stated by Hull (1943). This is an erroneous assumption, and because of this, many of Reid's criticisms are baseless. Let us therefore take Reid's points in order as he presents them: he first described I_R as being 'in the nature of a negative drive akin to tissue injury, fatigue or pain'. This is the disproven notion of I_R being a peripheral phenomenon. It is now regarded by most research workers as being centrally generated. Because of several omissions on the part of Reid his argument that he does not see how learning can take place seems real enough. However, he omits the Kimble (1949) drive/inhibition tolerance hypothesis, i.e. the greater the primary drive the greater the strength of I_R necessary before it becomes supraliminal and Eysenck's (1956) notion of the extinction of conditioned inhibition following a rest pause during which time reactive inhibition dissipates and becomes subliminal. There is no argument therefore against the possibility of learning if the positive habit is supraliminal and the negative drive and habit is kept to a subliminal level.

Reid links my name with Gleitman *et al.* (1954). I wish to point out that I was in serious disagreement with these workers, as can be seen from my article (Kendrick 1958a). It is not so unusual to find the deduction confirmed. From personal contact with Russian and Polish research workers the impression has been gained that this form of extinction has to be guarded against. Razran (1956a) has discussed this point quite fully. I have already discussed the probability of sI_R being equal to sH_R , and so no more need be said about this point. Concerning the learning equation, Gwynne Jones has taken up this criticism in his paper.

In conclusion, I still find that the concept of antagonistic habits explains my results adequately.

IV. INDIVIDUAL DIFFERENCES IN INHIBITORY POTENTIAL

By H. GWYNNE JONES

*Institute of Psychiatry, Maudsley Hospital, University of London**

Much of the earlier psychological research at the Maudsley Hospital was concerned with the definition and measurement of descriptive, static dimensions of personality, in particular *extraversion-introversion* and *neuroticism*. Eysenck's main interest lay, however, in the genetic and dynamic aspects of personality, in process rather than structure and, more recently, he has outlined a tentative theory to explain the genesis of the personality types related to these dimensions and of the two broad categories of neurosis, *hysteria* and *dysthymia* (Eysenck, 1957). Such a dynamic theory needs to refer not only to individual differences but also to general laws. As the development and modification of behaviour was to be explained, the generalizations derived from experiments in learning were considered most apt.

Any contemporary theory attempting to integrate knowledge concerning individual differences, genetic personality studies and the general laws of learning is very unlikely to be adequate to the entire range of data it needs to encompass. It is, however, possible to set up a conjectural theory in rough accord with a selection of the main facts and to state its postulates so as to be empirically testable. It is, of course, essential that the theory be tested vigorously over a wide range of its implications and, almost certainly, it will need progressively to be modified in the light of the findings. Even if findings are so discrepant that the original theory needs to be abandoned, the researches indicated in the validation process are likely to yield new data directly relevant to the original problems.

The empirical data concerning learning have been interpreted from many theoretical viewpoints and, for his attempt at integration, Eysenck adopted Hull's theory. Hull's system has been much criticized, but it had many advantages for this particular task. More than any other learning theorist, Hull attempted to describe a wide range of complex human behaviour as well as the laboratory learning of experimental animals. His postulates and assumptions are fairly clearly stated so that predictions may be made on to new situations for which previous experimental results are not available. A further advantage lay in the fact that important related researches, such as those of Spence and his colleagues at Iowa, derive from neo-Hullian theory. Perhaps most important, however, was Hull's employment of intervening variables to which individual differences could be related. It is difficult to postulate individual differences between empty organisms except in terms of their shape as modified by their previous history. There was evidence that many of the individual differences in question are of a genetic and constitutional nature.

Hull's concepts of inhibition have been particularly criticized, not least by Reid within this Symposium and elsewhere (Reid, 1958). Many of his points are valid:

* The opinions expressed in this paper are the author's own. They should not necessarily be considered as representing the point of view of Professor Eysenck.—H.G.J.

others, perhaps, display more wit than wisdom. The formula he quotes does not radically alter Hull's theory but follows logically from Hull's definitions of his variables. The state of affairs which Reid describes as absurd would occur rarely in practice but can exist in special experimental conditions. Pavlov established such conditions and reports 'ultraparadoxical' effects consistent with the implications of the formula. Perhaps 'Hullian Solitaire' can be a rewarding scientific exercise as well as an academic game! The difference between the use of plus and minus signs in this formula is precisely the difference between the addition of a negative quantity and the subtraction of the same but positive quantity.

Reid (1958) lays great stress on the role of reinforcers as stimuli and has attempted to reinterpret inhibitory phenomena in terms of changes in the stimulus situation. His analyses of disinhibition, speedy recovery when reinforcement is reinstated after extinction, and warm-up phenomena are very convincing. A very similar explanation of long-term warm-up has been incorporated into Eysenck's theory (Eysenck, 1956).

Extinction and reminiscence, i.e. partial spontaneous recovery from the end of one period of massed practice to the beginning of the next, are, however, more central to the theories under discussion. Reid interprets these, too, in terms of changing stimulus conditions. For him, extinction is a form of discrimination learning. If so, it is difficult to see why those who reach a criterion most rapidly during positive learning tend to extinguish least rapidly. In 'inhibition with reinforcement' situations, as described by Kendrick earlier in this Symposium, the stimuli do not change in the way required for Reid's process to operate and 'free reinforcement' at the end of the procedure fails to re-instate the original behaviour. Similar, more or less complete inhibition of a voluntary response may be achieved with human subjects (Jones, 1960*a*).

To claim usefulness for Hullian theory does not imply a claim that it cannot be improved. Indeed, once the learning theory was incorporated into the general theory, empirical tests of the latter applied as much to its learning theory components as to any other. Several difficulties of this nature have arisen: one example, relevant to this symposium, concerns an apparent disproof by Eysenck of the Hull-Lepley explanation of the bowing of serial position learning curves in terms of specific intra-list inhibition. By an experiment* which confirmed earlier work by Mitchell, it was shown that if a list of nonsense syllables was learned continuously—the inter-list and intra-list intervals being equal—the bowing effect did not differ from that produced by normal inter-list spacing. This result is also inconsistent with 'consolidation' explanations of the bowing. One intriguing subsidiary finding was that, with massed practice, some subjects 'lost their place' in the list and believed that one of the middle syllables was in fact the first. Even so, these subjects produced normal bowed curves. This suggests that the bowing is determined during the first presentation of a list.

Eysenck's (1957) general theory was expressed as a series of postulates. Of these, two are of central importance. The first, called the *Postulate of Individual Differences*, runs as follows:

Human beings differ with respect to the speed with which excitation and inhibition are produced, the strength of the excitation and inhibition produced, and the speed with which inhibition is dissipated.

* The majority of experiments reviewed in this paper derive from the current programme of research at the Institute of Psychiatry. Detailed reports will be published in due course.

The second, the *Typological Postulate*, runs roughly as follows:

Individuals in whom excitatory potential is generated weakly and slowly are thereby predisposed to develop extraverted patterns of behaviour and to develop hysterical-psychopathic disorders in cases of neurotic breakdown; individuals in whom excitatory potential is generated strongly and quickly are thereby predisposed to develop introverted patterns of behaviour and to develop dysthymic disorders in cases of neurotic breakdown. Similarly, individuals in whom reactive inhibition is generated quickly and strongly and dissipated slowly are thereby predisposed to develop extraverted patterns of behaviour and hysterical-psychopathic disorder in neurotic breakdown; individuals in whom reactive inhibition is generated slowly and weakly and dissipated quickly are thereby predisposed to develop introverted patterns of behaviour and dysthymic disorder in neurotic breakdown.

In these postulates Eysenck refers to excitatory and inhibitory potentials: an earlier version (Eysenck, 1955*a*) referred only to reactive inhibition. From his discussion of the postulates, he appears to consider excitation and inhibition in terms of a single bi-polar dimension. From the design of later experiments, however, it is clear that excitatory potential is to be measured in terms of *habit strength*, the sH_R component of Hull's equation, while inhibitory potential is to be measured in terms of *reactive inhibition* (I_R). Logically then these two dimensions may be independent and it is possible for an individual to generate both potentials equally strongly.

Although he makes no formal postulate, Eysenck considers the personality dimension of *neuroticism* to reflect individual differences in arousal of drive (D). This may relate particularly to avoidance drive in threatening situations.

Again without formal statement of a postulate, Eysenck stresses similarities between the processes of reactive inhibition and perceptual satiation. He suggests that individual differences in satiability and inhibitability are similarly related to personality differences.

The writer (Jones, 1960*b*) has been particularly interested in examining the finding of Spence (1956) and others concerning relationships between drive and performance and assessing their consistency with Eysenck's postulates. An extensive review of the literature yielded no wide generalizations until, in addition to the variables already mentioned, consideration was given to the stressfulness of the experimental situation in terms of the difficulty of the task, its pacing, and external stressors. The latter may be psychological in nature, such as reports of failure, or physical stressors such as electric shock. There are also strong indications of the general applicability of the Yerkes-Dodson principle which may be stated in terms of two laws:

- (1) 'Efficiency of learning is a curvilinear function of drive strength, some intermediate level of drive being optimal.'
- (2) 'Optimal drive strength is an inverse function of the difficulty of a learning task.'

When these additional factors were taken into consideration, the findings, with a few exceptions, were in general agreement with Eysenck's theory. However, from the nature of the personality tests used, this applies more to his hypotheses concerning drive and *neuroticism* than to those concerning inhibition and *extraversion*.

By no means the least advantage of an integrative theory of the type described is the frame of reference it provides for the clinical psychologist who has to attempt the description and, to some extent, the modification of the abnormalities of psychiatric patients (Jones, 1960*a*). An inhibitory phenomenon, not referred to in the postulates,

but particularly relevant to clinical psychology is that of reciprocal inhibition. If an acceptable response, incompatible with an abnormal response but linked with the same stimulus, can once be established, it may be progressively reinforced until it becomes prepotent in life situations evoking the symptoms (Wolpe, 1958; Jones, 1960b).

Much of the recent research in Professor Eysenck's department has been concerned with the testing of his inhibition/excitation postulates. The relevant experiments may be classified into four main categories.

1. *Conditioning*. Franks (1956) reported differences between *hysterics* and *dysthymics* in the acquisition and extinction of conditioned eye-blink responses and an appreciable negative correlation between conditioning and extraversion within groups of psychiatric and normal subjects. Later studies by Willett and others have produced correlations which, though consistently negative, have been frequently insignificant and always smaller than those reported by Franks. From all the data now available the best estimate of the size of the correlation would be of the order of -0.20 .

Willett employed three other conditioning techniques. Salivary conditioning scores correlated positively and significantly but only slightly with some measures of neuroticism, and negatively but insignificantly with extraversion. So-called verbal conditioning scores derived from a situation in which, over a long series of trials, the subjects were required to predict the direction (right or left) of a light following a red warning light. These scores correlated significantly with neuroticism measures but insignificantly with extraversion. Spatial conditioning scores reflected the influence of the location of a light signal on judgements concerning the location of a hidden source of a tone. One such measure was significantly but slightly correlated with extraversion and scores on a 'lie' scale. All these conditioning scores were of low reliability.

In so far as all four techniques may be described as conditioning and their reliabilities are adequate, the intercorrelations between scores should reflect the general factor of conditionability which would be predicted from Eysenck's theory. Measures from the salivary, spatial, and verbal procedures tend to intercorrelate positively and significantly. The coefficients, however, are only of the order of 0.20 to 0.30 and any common factor is unlikely to account for more than a small proportion of the variance. The lack of relationship between eye-blink conditioning and the other measures may reflect the distinction between avoidance and non-avoidance conditioning.

Eysenck (1959) found that verbal conditioning of the now familiar 'Hm-mm' type discriminated significantly on a one-tail test between groups of neurotic patients selected in terms of extreme scores on the extraversion (*E*) scale of the *Maudsley Personality Inventory* (*M.P.I.*) Initial differences between the responses of the two groups were, however, more striking.

In general, Eysenck's theory, at least in so far as it refers to excitation as distinct from inhibition, gains little support from recent work on conditioning. Relationships with neuroticism are observed at least as frequently as those with extraversion and the latter are seldom strong.

II. *Reminiscence*. More relevant to the inhibition components of the theory are studies of the reminiscence phenomenon. Several experiments by Eysenck, Star and others were designed to test the prediction that reminiscence would be greater for extraverts than introverts. This prediction was consistently and significantly verified for the first reminiscence score during serial massed practice of rotary pursuit. The correlation reduces to insignificance for later reminiscence scores in the series. Significant correlations between reminiscence and neuroticism were also observed but these were positive in some conditions and negative in others. The latter finding may be related to the Yerkes-Dodson effect mentioned earlier (Jones, 1960*b*).

One intriguing finding derives from a well-designed experiment by Treadwell which produced highly significant results consistent with the theory. She employed a track-tracing task and the *Thinking Extraversion* (*T*) scale of the Minnesota *T.S.E.* scale of extraversion. All available evidence indicates that the Minnesota *T* scale is uncorrelated with the *E* scale of the *M.P.I.*

In this area, then, the original hypotheses are still viable and further research is well justified.

III. *Perception*. The first test of the hypothesis that satiation effects are correlated with extraversion involved the kinaesthetic figural after-effect phenomenon described by Klein and Krech (Eysenck, 1955*b*). Relationships of the type predicted have been obtained by Eysenck, Nichols, and Broadbent. There are a number of inconsistencies between detailed findings and again there is evidence that neuroticism also correlates with satiation scores. Completely invalidating Eysenck's theoretical explanation, however, Broadbent elegantly showed that the direction of group differences could be reversed by reversing the orientation of the apparatus.

Nichols applied a battery of tests, all considered to measure satiation and figural after-effects, in the kinaesthetic, visual, and auditory modalities. There was no evidence for a general factor of satiability and the individual results gave little support to the theory.

More recently, Holland and others have investigated relationships between questionnaire measures of personality and various perceptual functions including auditory flutter fusion, foveal and peripheral visual flicker fusion, contour and brightness figural after-effects, visual field perimetry, and negative after-effects of the Archimedes spiral. The spiral after-effect was shown to be a reliable phenomenon which was frequently and significantly related to extraversion but only insignificant or very low correlations were observed for the remainder.

Despite possible explanations in terms of the many practical and methodological difficulties of research in this area, it is only justified to conclude that Eysenck's hypotheses concerning satiation have so far failed to be verified.

IV. *Drug studies*. Many experiments have been conducted to verify a postulate which runs as follows:

Depressant drugs increase cortical inhibition, decrease cortical excitation, and thereby produce extraverted behaviour patterns. Stimulant drugs decrease cortical inhibition, increase cortical excitation, and thereby produce introverted behaviour patterns.

Various deductions from this postulate have been experimentally verified at adequate levels of significance by experiments in which stimulant and depressant

drugs have been shown to have the predicted effects on a variety of measures of conditioning, learning, perceptual, physiological, and behavioural phenomena.

Although clearly in accord with the predictions made, these results do not, in the writer's view, unequivocally verify the postulate. The latter refers only to excitation and inhibition whereas the drugs used are quite likely to affect the drive states of the subjects. From Hullian theory, changes in drive would be expected to affect performance in a similar manner to changes in habit strength or inhibitory potential. In terms of the theory as a whole, it is inconsistent that the drug studies involving certain perceptual tasks tend to support the drug postulate whereas the predicted relationships between the same perceptual variables and psychometric measures of extraversion were not observed or account for very little of the variance in performance.

Finally, then, it seems fair to conclude that the original theory is likely to require considerable modification. Statements concerning perceptual phenomena may need to be excluded and the individual difference and typological postulates may need to be restricted by referring only to inhibitory and not to excitatory potentials. In other ways the theory needs to be expanded so as to refer, for example, to interaction factors between neuroticism and extraversion/introversion and to the effects of stress. This state of flux, however, merely reflects its status and functions as a scientific theory and not as a dogma.

Even in the areas where experimental support was not forthcoming, much research has been stimulated and data obtained relevant to the problems of personality research and abnormal psychology. More researches are in progress or are planned.

V. INHIBITION—PAVLOV, HULL, EYSENCK

BY R. L. REID

Department of Psychology, University of Aberdeen

1. PAVLOV

Pavlov (1927) considered himself a physiologist, using the conditioning experiment as a new tool for exploring the action of the cerebral hemispheres. This does not make him easy to understand.

He described the action of the cortex (behaviour) in terms of excitation and inhibition. The two processes have properties in common, so he thought. They irradiate and concentrate according to the same laws and each limits the action of the other. Pavlov interpreted conditioning as a linking of stimuli (signalization) rather than of stimulus and response. Accordingly, he located the excitatory and inhibitory processes in the cortical cells of the sensory analysers.

Two main types of inhibition were distinguished. The first, external inhibition, observed as a straightforward interference effect, need not concern us. The second, internal inhibition, is more relevant.

Pavlov observed that as a result of giving extinction trials a conditioned stimulus could become inhibitory rather than merely neutral, i.e. the stimulus could be shown to weaken the effects of other active conditioned stimuli if presented before them. This, together, with observations of what he called 'spontaneous recovery' and 'disinhibition' led him to regard extinction as an inhibitory effect, not a simple return to the neutral state prevailing before reinforcement.

He suggested that the aggregate duration of excitation produced by the c.s. acting alone might be the factor determining the accumulation of inhibition. On reinforcement trials periods of excitation are cut short, so producing less inhibition than on extinction trials. Below a minimum, inhibition, he thought, might not accumulate from trial to trial. However, he also stated that c.r.'s inevitably weaken even if uniformly reinforced, so that it is not clear what he believed.

There is a further difficulty. Pavlov attributed the inhibitory effects of repeated stimuli to functional exhaustion of the active cortical cells. But since he also said that hand in hand with exhaustion goes recovery, this does not seem to explain chronic extinction effects of any type. Konorski (1948) mentioned the difficulty and said that in his later years Pavlov believed that permanent inhibitory states might occur through reductions in the 'capability' of cortical cells. (Any cell if stimulated very strongly reacts with inhibition rather than excitation, and a reduction in capability would result in inhibition at lower intensities, perhaps even down to zero intensity.) Pavlov seems not to have been sufficiently satisfied with this idea to put it in writing. Perhaps as a true physiologist he felt more at home when dealing with inhibition as a process rather than as a permanent disposition.

Pavlov was a pioneer with a flair for making discoveries and an intense interest in facts. His theories, like those of the Ethologists in a different field, provide a convenient way of talking about behaviour in certain limited situations. At this level an

inhibition theory of extinction may be convenient, but attempts to develop it by, for example, Konorski (1948) and Hull (1943) have not proved satisfactory. The notion that an extinguished response is somehow 'there all the time', but inhibited, may provide more problems than it solves.

II. HULL

Hull developed Pavlov's view of extinction as an inhibitory process but it should be noticed that Hull's general approach was very different from Pavlov's. The Hullian system is based on a version of the Law of Effect. It employs some Pavlovian terms, but it applies more directly to instrumental learning than to classical conditioning.

When learning, a Hullian animal builds up a habit strength, sH_R , by establishing connexions between receptor and effector organs. Habit strengths are mere dispositions to respond, and actual responses take place only under suitable conditions of motivation, D . sH_R multiplied by D gives reaction potential sE_R . Since bigger and better incentives increase performance, a variable, K , incentive motivation, is introduced, again as a multiplier. With some other complications that are not relevant here, performance is predicted from sE_R after allowing a subtraction due to the inhibitory aggregate or potential I_R .

This inhibitory aggregate comprises reactive inhibition, I_R , a temporary after-effect of all responses 'in the nature of a negative drive akin to tissue injury, fatigue or pain', and conditioned inhibition sI_R , a permanent disposition to stop performing a response, set up in the same way as sH_R through the reinforcement provided by ceasing to respond.

A difficulty becomes apparent. Reactive inhibition occurs during reinforcement as well as extinction trials. If ceasing to respond can be conditioned in the same way as an ordinary habit it is difficult to see how learning can take place at all. Ceasing to respond is reinforced through a reduction in I_R and also through the reinforcement provided by the food reward or whatever is used in the experiment. Performing the response is reinforced by the same events but ceasing to respond is nearer in time to both processes of need reduction and so should gain a larger increment of strength. Hull first saw this difficulty in 1943. He suggested that while the response was being performed secondary reinforcement would occur, e.g. through the click of the food mechanism in a Skinner box. But this is no solution. A learning process must start before such stimuli can even acquire their secondary reinforcing powers.

If we now shelve this problem, as Hull did, and assume that somehow sH_R can build up more quickly than sI_R we can 'predict', as they say, that learning may occur. But another difficulty now appears. Both the excitatory and the inhibitory dispositions grow exponentially and when they reach their asymptotes further reinforcement or extinction trials can have no effect on reaction potential. Gleitman *et al.* (1954) and Kendrick (1958*a*) deduced from Hull's theory that responses would inevitably extinguish in spite of continued reinforcement. Gleitman *et al.* regarded this as a serious flaw, while Kendrick appeared to be satisfied and sought out cases where the deduction was confirmed. Without doubt there are in fact cases where responses diminish despite continued reinforcement, but these are exceptions. Pavlov's dogs did not cease to salivate when shown meat, and human beings do not appear to extinguish walking, speaking, toilet training, etc. It is good to have a theory to explain

the occasional blueness of the moon provided that you have another one accounting for its normal yellowness.

In passing it should be noticed that Gleitman *et al.* and Kendrick make the assumption that the asymptote of sI_R is higher than that of sH_R . The opposite might just as well be true, and in such cases responses would become permanently fixated and impossible ever to extinguish. Thus Hull's theory appears to be able to 'predict' anything but the ordinary results of experiments on learning and extinction.*

The difficulties that have been discussed are in a sense latent in Pavlov's theory. Had Pavlov been seriously concerned with explaining permanent extinction effects he would have met with some of them at least. Konorski (1948) provided a way out by postulating that the occurrence of inhibition could create new possibilities of excitation but he did not develop the notion beyond the *ad hoc* stage.

Looking now at some of the detail of Hull's theory of extinction, a different kind of difficulty becomes apparent. I_R is described as a primary negative drive generated by responding, and sI_R as a conditioned disposition to stop making a specific response. If I_R and sI_R are merely added as Hull suggests, how is sI_R activated? Unless it is activated there can be no permanent extinction. Plain multiplication of sI_R by I_R is not a solution because then the permanent extinction component sI_R could not be activated until some I_R had been produced by the occurrence of at least one response.

Gwynne Jones (1958) suggested that the way out was to subtract sI_R from sH_R and I_R from D , giving the formula

$$s\bar{E}_R = f(D - I_R) \times f(sH_R - sI_R).$$

But this is absurd. Let sI_R be greater than sH_R and I_R will be greater than D . A positive reaction potential will then result through the multiplication of two minus quantities, i.e. other things equal, the greater the drive for not responding and the stronger the habit of not responding, the more likely it would be that a response would be made. I have indulged in this sort of multiplication only because Gwynne Jones does.†

* Kendrick seems to believe that the asymptotes of sI_R and sH_R are necessarily equal, and that a learned response will no longer be performed when equality obtains. But equality would mean that the organism would neither *perform* the response in question nor would it *not perform* the response. It is tempting to suppose that, faced with this dilemma, it would go the way of the Cheshire Cat. However, Hull employs the notion of oscillation which, while not avoiding the logical difficulty that has been mentioned, would lead to the prediction of a 50 % response rather than extinction. Kendrick says: 'The asymptote of sI_R like that of sH_R will depend upon the activating drive strength. If we accept the Jones formula for learning then sI_R must equal sH_R in time because both are activated by the same drive level.' But the formula in question concerns the calculation of sE_R and not the growth of sI_R or sH_R , and even if they were both determined by the same prevailing drive strength (contra Hull) there are other determinants of their asymptotes. Only an *ad hoc* assumption could prevent sI_R from being as likely as not to reach a lower level than sH_R .

† According to Gwynne Jones this would rarely happen and is represented by Pavlov's ultraparadoxical phase. But this phase as specified by Pavlov was a transient phenomenon. The formula predicts something different. Whenever the two resultants in brackets in the equation go negative, the response in question will be performed. The effect of performance will be to increase I_R and sI_R resulting in ever-increasing sE_R and so stronger responding, the animal working itself up by positive feedback into a fury of sE_R until exhaustion supervened. To take up the other point, I do not see why the condition should be rare if the formula is accepted as correct, but in any case the formula produces odd predictions when only one of the resultants is negative and this condition would certainly be common enough since negative values of sE_R are required to account for extinction. When I_R is greater than D , sE_R will be *positively* related to sI_R , and *negatively* related to sH_R . Similarly, when sI_R is greater than sH_R , sE_R will be *positively* related to I_R and *negatively* related to D . If true, this is a real psychological discovery unsuspected by common sense.

Hull did not deal with the difficulty but it is fairly clear that a consistently Hullian animal would have three inhibitory components to add up and multiply. IR_1 as a primary negative drive would give rise to a conditioned negative drive IR_2 and these would summate to give a total negative drive. sIR would be formed through reduction of this total drive. A situation in which conditioning and extinction had taken place would set off IR_2 and so sIR would be activated even in the complete absence of IR_1 . Thus the difficulty can be got over, but improvements of this sort are about as useful as rearranging the flowers in a building that is on fire. The whole theory needs drastic revision.

I have said nothing so far about K . The advanced Hullian plays K as a sort of Joker. Incentive motivation is a multiplier co-ordinate with D and sH_R and the use of K involves a learning theory within a theory. The effects of changes in incentive are registered very quickly—in the example Hull gave (1952) at 80 % of maximum in one trial. Thus within two extinction trials the value of K would go down by 96 %. By itself K can account for more extinction than has generally been observed.

Without mentioning fractional antedating goal responses, enough has been said to indicate what may have prompted Professor Drever to refer to Hull's model as an Emmett engine. Alternatively, the theory can be thought of as an academic game with several varieties, including Hullmanship and Hullian Solitaire. We are all familiar with the first; the second is played by formulating an experimental procedure and 'predicting' all physically possible outcomes.

III. EYSENCK

In 1955, Eysenck put forward the postulate: 'Human beings differ with respect to the speed with which reactive inhibition is produced, the strength of the reactive inhibition produced, and the speed with which reactive inhibition is dissipated....' Three variables were mentioned but there was no attempt to distinguish between them and it seemed that a high degree of correlation between the three was implied. A second postulate asserted: 'Individuals in whom reactive inhibition is generated quickly, in whom strong reactive inhibitions are generated, and in whom reactive inhibition is dissipated slowly are thereby predisposed to develop extraverted patterns of behaviour and to develop hysterico-psychopathic disorders in cases of neurotic breakdown....' The converse was stated for introversion and dysthymic disorders.

In 1955, Eysenck depended largely on the results of an experiment by Franks on eyeblink conditioning. Again, in 1957, he wrote: 'We have relied largely on the conditioned eyeblink for evidence in respect of deductions made from our typological postulate.' It seems worth looking closely at this experiment.

Franks (1956) conditioned and extinguished the eyeblink response of a group of normal subjects and two groups of patients classed as dysthymic or as hysteric, in Eysenck's sense. He found that the dysthymics gave more conditioned responses during both reinforcement and conditioning trials than the hysterics, the normal group giving an intermediate performance. The differences between the dysthymic group and the other two groups were significant.

This experiment seems to offer little support to Eysenck's postulates, although at the time he appeared to think it did.

(1) Considering the extinction data, if, as asserted, the groups reached different levels of conditioning at the end of the fixed number of reinforcements given to all subjects, then regardless of Eysenck's postulate the highest performing group would be expected to give more responses.

(2) Considering rates of conditioning, the average interval between reinforcements was 40 sec. This is sufficiently long to produce near-optimal conditioning. It must then be long enough to allow reactive inhibition to dissipate between trials to an insignificant level, in which case any differences in rates of conditioning must be referred to some other factor. Eysenck (1957) seemed to agree, since he then wrote: 'General theoretical considerations make it appear unlikely that in the present experiment inhibitory potentials would assume a very prominent place....'

If, as it might now appear, the results of experiments like Franks's on rates of conditioning are to be attributed to differences in excitation rather than inhibition, the connexion between these results and the results of other experiments grouped together by Eysenck, e.g. experiments on perceptual satiation or reminiscence, can be made only via a theory that people who acquire sH_R 's quickly will also develop inhibition slowly. This does not seem unreasonable. From a commonsense point of view excitation and inhibition may be related in this way. But we are not dealing with common sense. Within the context of Eysenck's later theory any relationship which might exist between, for example, the rapid setting up of conditioned blinking and the slow development of a figural after-effect is arbitrary and unexplained. The verbal link formerly provided by the work 'inhibition' has been broken. Eysenck uses the blanket term 'excitation-inhibition balance' but this does not seem to contribute anything new since in the theory excitation and inhibition still produce their separate potentials and do not interact.

In his 1957 version of the postulates significant additions were made. Human beings were now said to differ in the speed with which both excitation and inhibition are produced, in the strength of the excitation and inhibition produced, and the speed with which inhibition is dissipated. There were now five dimensions of variation, with inter-relationships that were implied although not stated or tested.

Also in 1957, Eysenck added a sixth possible variable to the five that had already been said to differentiate introverts from extraverts. He suggested that introverts and extraverts might differ in the asymptote which inhibition must approach before causing involuntary rest pauses. The involuntary rest pause idea had been put forward by Kimble (1949), who suggested that these pauses would occur whenever negative drive had accumulated up to the level of positive drive prevailing at the time. On the occurrence of an *IRP* the reinforcement due to resting would generate sI_R .

If sI_R is conceived to build up only when *IRP*'s occur, some of the difficulties in the Hullian theory are removed and we begin to be able to 'predict' that learning will sometimes occur. But the new position is just as unsatisfactory as the old. There is no way of explaining how sI_R could develop in experiments such as Franks's. Experimental trials in this case lasted only a fraction of a second. Where are the involuntary rest pauses? How could extinction ever take place?

With a system as complicated as Hull's a danger is that improving its explanation of *X* will impair its explanation of *Y*. Eysenck (1957) seems to be unaware of this and so oblivious of the system as a whole that in a discussion of involuntary rest

pauses he was prepared to write: 'During the period...while $D-$ is growing but has not yet reached the level of $D+$, no involuntary rest pauses occur to reinforce sI_R and, in accordance with the laws of learning, sI_R being un-reinforced, undergoes extinction.' If this were so, a few extinction trials daily would permanently revive an extinguished response, but we should not take the statement seriously. Eysenck is cheating at Hull. Only *responses* can be reinforced or extinguished and sI_R is not a response. The extinction of a response is accounted for by the growth of sI_R . How can sI_R itself extinguish? At this point it becomes clear that a complete new theory is needed.

Finally the connexion between inhibition and personality should be examined. Eysenck suggested that 'socialization' distinguishes introverts from extraverts. 'Socialization' occurs through the development of a superego or conscience or set of acquired emotional drives, mainly anxiety. This could be called the Freud-Eysenck theory. If introverts condition more easily than extraverts they should become more 'socialized', i.e. socially anxious. However, it can be inferred from Eysenck's theory that introverts should not only learn to be more 'socialized' than extraverts. They should be better at all learned activities. They should talk earlier: they should perform better at school subjects: they should learn faster to swim, skate, speak French, and play the violin. Do they? The theory seems to account for more than has been observed.

No one could be out of sympathy with Eysenck's aims or refuse to admire some of his *tours de force*, but it might be wiser to wait at least until some learning theory successfully accounts for the behaviour of a rat in a box before extending it to explain differences in human personality. According to Eysenck, learning theory is the best developed area of experimental psychology. There is another point of view indicated by the complaint that learning theory is taking a long time to catch up with common sense. No theory, even when used properly, is in fact good enough to do what Eysenck or any other personality theorist would like it to do, and in talking about personality it might be better to use everyday notions about learning until some degree of agreement about the basic processes has been reached.

IV. CONCLUSION

Any account of learning must take into account the interference of incompatible responses. In a given situation a response occurs not because it has a fixed absolute strength but because it is stronger than other incompatible behaviour. Hull (1952) postulated: 'When the net reaction potentials (sE_R) to two or more incompatible reactions (R) occur in an organism at the same instant, each in a magnitude greater than sL_R , only that reaction whose momentary potential (sE_R) is greatest will be evoked.' The extinction of one response does not lead to a behavioural vacuum but to the performance of some other response. For example, in an experiment by Hull (1934), rats were first reinforced after running 20 ft. down an alley; then reinforcement was given only at 40 ft., and finally the rats were extinguished. During extinction trials the habit of halting at the 20 ft. point made its reappearance.

It scarcely seems necessary for Hull to have supposed that in extinction a special kind of 'not responding' would have to be conditioned. Detailed accounts of

behaviour in extinction are rare. Usually, behaviour has simply been classified as 'responding' and 'not responding'. By definition any weakening of the observed 'response' could be counted as a strengthening of the 'not-response'. But there are many possible ways of not responding. At the moment, I am not playing the violin or standing on my head, but that does not mean that I am doing nothing or that I am inhibiting such actions. However, psychologists have generally arranged experimental conditions so that in fact when an animal is not responding it is most likely doing nothing very much, simply because there is nothing much for it to do in, for example, a Pavlovian harness or a familiar Skinner box. Without this artifact pure interference theories lacking any concept of the sI_R type might have been more commonly held.*

Attempts to assess separately the absolute strengths of inhibitory and excitatory potentials inevitably lead to the Hullian dilemmas. A more workable form of inhibition theory would allow inhibition and excitation to interact as processes, adding or subtracting net amounts to or from probability that a response will be made in a given situation. But this form of theory, however plausible, must give way on the grounds of parsimony to a single-process theory that can account for the same facts. Even on its home ground, for example in explaining spontaneous recovery, the inhibition theory is unnecessarily complicated. Why should we expect that what has been learned at the end of one experimental session will carry over completely to the rather different stimulus conditions that prevail at the beginning of the next? But those who play the game of theorizing for its own sake may continue to juggle with two inhibitory processes. Perhaps they will be grateful to me for providing them with a third, earlier in this paper.

* Gwynne Jones states an objection that has often been made against interference theories—that it is difficult to see why those who reach a criterion most rapidly during positive learning tend to extinguish least rapidly. I cannot accept that this generalization has in fact been established. Even if it were established the relationship would certainly not be difficult to account for without postulating sI_R . For example, relatively active rats in a lever pressing situation might be expected to condition more quickly and extinguish less quickly than relatively inert rats.

BIBLIOGRAPHY

- AIKEN, E. G. (1957). The effort variable in acquisition, extinction, and spontaneous recovery of an instrumental response. *J. Exp. Psychol.* **53**, 47-51.
- ARIETI, S. (1951). Special logic of schizophrenic and other types of autistic thought. Reprinted in E. L. Hartley, et al.: *Outside Readings in Psychology*. 1951. New York: Crowell.
- ARIETI, S. (1955). *Interpretation of Schizophrenia*. New York: Brunner.
- BARNETT, S. A. (1958). Exploratory behaviour. *Brit. J. Psychol.* **49**, 289-310.
- BEACH, H. D. (1957). Effect of morphine on the exploratory drive. *Canad. J. Psychol.* **11**, 237-44.
- BERLYNE, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.* **41**, 68-80.
- BERLYNE, D. E. (1951). Attention to change. *Brit. J. Psychol.* **42**, 269-78.
- BERLYNE, D. E. (1955). The arousal and satiation of perceptual curiosity in the rat. *J. Comp. Physiol. Psychol.* **48**, 238-46.
- BINDRA, D. (1959). *Motivation: a Systematic Reinterpretation*. New York: Ronald Press.
- BRADLEY, P. (1957). The central action of certain drugs in relation to the reticular formation of the brain. In *Reticular Formation of the Brain*. Henry Ford Hospital International Symposium. London: Churchill.
- BROADBENT, D. E. (1953). Classical conditioning and human watch keeping. *Psychol. Rev.* **60**, 331-9.
- BROADBENT, D. E. (1959). Human perception and animal learning. To appear in a volume edited by O. L. Zangwill and W. H. Thorpe.
- BROWN, J. S. (1953). Comments on Professor Harlow's paper in *Current Theory and Research in Motivation*. University of Nebraska Press.
- BURES, J., BURESOVA, O. & ZAHAROVA, A. (1958). Conditional reflexes and Leao's spreading cortical depression. *J. Comp. Physiol. Psychol.* **51**, 263-8.
- BUTLER, R. A. (1953). Discrimination learning by Rhesus monkeys to visual-exploration motivation. *J. Comp. Physiol. Psychol.* **46**, 95-8.
- DOLLARD, J. & MILLER, N. E. (1950). *Personality and Psychotherapy*. New York: McGraw Hill.
- EISMAN, E. (1956). An investigation of the parameters defining drive (D). *J. Exp. Psychol.* **52**, 85-9.
- ELLIS, D. S. (1953). Inhibition theory and the effort variable. *Psychol. Rev.* **60**, 383-92.
- ENGLISH, H. B. & ENGLISH, A. C. (1958). *A Comprehensive Dictionary of Psychological and Psychoanalytical Terms*. New York: Longmans.
- ESTES, W. K. & SCHOEFFLER, M. S. (1955). Analysis of variables influencing alteration after forced trials. *J. Comp. Physiol. Psychol.* **48**, 357-62.
- ESTES, W. K. (1956). Learning. In *Annual Review of Psychology*, vol. 7, pp. 1-38.
- EYSENCK, H. J. (1955a). A dynamic theory of anxiety and hysteria. *J. Ment. Sci.* **101**, 28-51.
- EYSENCK, H. J. (1955b). Cortical inhibition, figural after-effect, and theory of personality. *J. Abnorm. Soc. Psychol.* **51**, 94-106.
- EYSENCK, J. H. (1956). 'Warm-up' in pursuit rotor learning as a function of the extinction of conditioned inhibition. *Acta Psychol.* **12**, 349-70.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London, Routledge and Kegan Paul.
- EYSENCK, H. J. (1959). Personality and verbal conditioning. *Psychol. Rep.* **5**, 570.
- FLUGEL, J. C. (1933). *A Hundred Years of Psychology*. London: Duckworth.
- FLUGEL, J. C. (1934). *Men and their Motives*. London: Kegan Paul.
- FLUGEL, J. C. (1955). *Studies in Feelings and Desire*. London: Duckworth.
- FRANKS, C. M. (1956). Conditioning and personality: a study of normal and neurotic subjects. *J. Abnorm. Soc. Psychol.* **52**, 143-50.
- FREUD, S. (1917). Mourning and melancholia. Reprinted in the *Standard Edition of the Complete Psychological Works of Sigmund Freud*. (Ed. J. Strachey,.) vol. xiv, 1957. London: Hogarth.
- FREUD, S. (1938). *The basic Writings of Sigmund Freud*. (Ed. A. A. Brill) New York: Modern Library.
- FULTON, J. F. (1949). *Physiology of the Nervous System*. 3rd ed. New York: Oxford University Press.

- GLANZER, M. (1953). Stimulus satiation: an explanation of spontaneous alternation and related phenomena. *Psychol. Rev.* **60**, 237-68.
- GLANZER, M. (1958). Curiosity, exploratory drive and stimulus satiation. *Psychol. Bull.* **55**, 302-15.
- GLEITMAN, H., NACHMIAS, J. & NEISSER, U. (1954). The S-R reinforcement theory of extinction. *Psychol. Rev.* **61**, 23-33.
- GLICKMAN, S. E. (1958). Effects of peripheral blindness on exploratory behaviour of the hooded rat. *Canad. J. Psychol.* **12**, no. 1, 45-51.
- GROSSLIGHT, J. H. & TICKNOR, W. (1953). Variability and reactive inhibition in the Med worm as a function of determined turning sequences. *J. Comp. Physiol. Psychol.* **46**, 35-8.
- HARLOW, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *J. Comp. Physiol. Psychol.* **43**, 289-94.
- HARLOW, H. F. (1953a). Mice, men, monkeys and motives. *Psychol. Rev.* **60**, 23-32.
- HARLOW, H. F. (1953b). Motivation as a factor in the acquisition of new responses. In *Current Theory and Research in Motivation*. University of Nebraska Press.
- HARLOW, H. F. (1959). Learning set and error factor theory. In S. Koch, (ed.), *Psychology: a Study of a Science*, vol. 2, pp. 492-537. New York: McGraw Hill.
- HARLOW, H. F. & HICKS, L. H. (1957). Discrimination learning theory: uniprocess versus duoprocess. *Psychol. Rev.* **64**, 104-9.
- HARRIS, J. (1943). Habitulatory response decrement in the intact organism. *Psychol. Bull.* **40**, 385-422.
- HAVELKA, J. (1956). Problem-seeking behaviour in rats. *Canad. J. Psychol.* **10**, 91-7.
- HEAD, H. & HOLMES, G. (1911). Sensory disturbances from cerebral lesions. *Brain*, **34**, 102.
- HEATHERS, G. L. (1940). The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.* **10**, 359-80.
- HEBB, D. O. (1949). *The Organization of Behaviour*. New York: Wiley.
- HEBB, D. O. & MAHUT, H. (1955). Motivation et recherche du changement perceptif chez le rat et chez l'homme. *J. Psychol. norm. Path.* **52**, 209-21.
- HEBB, D. O. & THOMPSON, W. R. (1954). The social significance of animal studies. In G. Lindzey (ed.) *Handbook of Social Psychology*. Cambridge: Addison-Wesley.
- HEBB, D. O. (1958). *A Textbook of Psychology*. London: Saunders.
- HERTZBERG, A. (1929). *The Psychology of Philosophers*. London: Kegan Paul.
- HILGARD, E. R. (1948). *Theories of Learning*. New York: Appleton-Century-Crofts.
- HILGARD, E. R. (1953). *Introduction to Psychology*. New York: Harcourt Brace.
- HOVLAND, C. I. (1936). 'Inhibition of reinforcement' and phenomena of experimental extinction. *Proc. Nat. Acad. Sci., Wash.*, **22**, 430-33.
- HULL, C. L. (1934). The rat's speed-of-locomotion gradient in the approach to food. *J. Comp. Psychol.* **17**, 393-422.
- HULL, C. L. (1943). *Principles of Behaviour*. New York: Appleton-Century-Crofts.
- HULL, C. L. (1951). *Essentials of Behaviour*. New Haven: Yale University Press.
- HULL, C. L. (1952). *A Behaviour System*. New Haven: Yale University Press.
- HUMPHREY, G. (1933). *The Nature of Learning*. London: Methuen.
- IWAHARA, S. (1957). Hull's concept of inhibition: a revision. *Psychol. Rep.* **3**, 9-10.
- JACKSON, J. HUGHLINGS (1931). *Selected Writings*, 2 vols. (ed. J. Taylor). London: Hodder and Stoughton.
- JONES, H. GWYNNE, (1958). The status of inhibition in Hull's system: a theoretical revision. *Psychol. Rev.* **65**, 179-82.
- JONES, H. GWYNNE (1960a). Continuation of Yates' treatment of a 'tiqueur.' In H. J. Eysenck, (ed.), *Behaviour Therapy and the Neuroses*. London: Pergamon.
- JONES, H. GWYNNE (1960b). Applied abnormal psychology. In H. J. Eysenck (ed.), *Handbook of Abnormal Psychology*. London: Pitman.
- KENDRICK, D. C. (1958a). Inhibition with reinforcement. *Bull. Brit. Psychol. Soc.* **34**, 4a.
- KENDRICK, D. C. (1958b). Inhibition with reinforcement (conditioned inhibition). *J. Exp. Psychol.* **56**, 313-18.
- KENDRICK, D. C. (1958c). *sI_R* and drive level: a reply to Keehn and Sabbagh. *Psychol. Rep.* **4**, 686.

- KENDRICK, D. C. (1959). An experimental investigation of inhibition with reinforcement in rats (conditioned inhibition). Unpublished Ph.D. thesis, University of London.
- KIMBLE, G. A. (1949). An experimental test of a two-factor theory of inhibition. *J. Exp. Psychol.* **39**, 15-23.
- KLÜVER, H. & BUCY, P. C. (1937, 1938, 1939). Quoted S. Arieti (1955), *Interpretation of Schizophrenia*. New York: R. Brunner.
- KONORSKI, J. (1948). *Conditioned Reflexes and Neuron Organization*. Cambridge University Press.
- LANDFRIED, D. E. & WIKE, E. L. (1956). A note on 'Experimental validation of conditioned inhibition'. *Psychol. Rep.* **2**, 485-8.
- LEPLEY, W. M. & RICE, G. E. (1952). Behaviour variability in paramecia as a function of guided act sequences. *J. Comp. Physiol. Psychol.* **45**, 283-6.
- LEUBA, C. (1955). Toward some integration of learning theories: the concept of optimal stimulation. *Psychol. Rep.* **1**, 27-33.
- McKELLAR, P. (1957). *Imagination and Thinking: a Psychological Analysis*. London: Cohen and West; New York: Basic Books.
- MONTGOMERY, K. C. (1952a). Exploratory behaviour and its relation to spontaneous alternation in a series of maze exposures. *J. Comp. Physiol. Psychol.* **45**, 50-7.
- MONTGOMERY, K. C. (1952b). A test of two explanations of spontaneous alternation. *J. Comp. Physiol. Psychol.* **45**, 287-93.
- MONTGOMERY, K. C. (1953). Exploratory behaviour as a function of 'similarity' of stimulus situations. *J. Comp. Physiol. Psychol.* **46**, 129-33.
- MYERS, A. K. & MILLER, N. E. (1954). Failure to find a learned drive based on hunger; evidence for learning motivated by 'Exploration'. *J. Comp. Physiol. Psychol.* **47**, 428-36.
- OSGOOD, C. E. (1953). *Method and Theory in Experimental Psychology*. New York: Oxford University Press.
- PAVLOV, I. P. (1927). *Conditioned Reflexes*. London: Oxford University Press.
- PAVLOV, I. P. (1955). *Selected Works*. (Ed. Kh. S. Koshtoyants.) Moscow: Foreign Languages Publishing House.
- PROSSER, C. L. & HUNTER, W. S. (1936). The extinction of startle responses and spinal reflexes in the white rat. *Amer. J. Physiol.* **117**, 609-18.
- RAZRAN, G. H. S. (1956a). Avoidant versus unavoidable conditions and partial reinforcements in Russian laboratories. *Amer. J. Physiol.* **69**, 127-9.
- RAZRAN, G. H. S. (1956b). Extinction re-examined and re-analysed: a new theory. *Psychol. Rev.* **63**, 39-52.
- REID, R. L. (1958). The role of the reinforcer as a stimulus. *Brit. J. Psychol.* **49**, 202-9.
- RICE, G. E. & LAWLESS, R. H. (1957). Behaviour variability and reactive inhibition in the maze behaviour of *Planaria dorotocephala*. *J. Comp. Physiol. Psychol.* **50**, 105-8.
- RICKMAN, J. (1957). *Selected Contributions to Psychoanalysis*. London: Hogarth.
- RIVERS, W. H. R. (1923). *Conflict and Dream*. London: Kegan Paul.
- RUSSELL, W. RITCHIE (1959). *Brain, Memory, Learning*. Oxford: Clarendon Press.
- SHARPLESS, S. (1954). Role of the reticular formation in habituation. Unpublished Ph.D. thesis. McGill University.
- SHARPLESS, D. E. & JASPER, H. (1956). Habituation of the arousal reaction. *Brain*, **79**, 655-80.
- SHERINGTON, C. S. (1906). *The Integrative Action of the Nervous System*. London: Constable.
- SILBERER, H. (1909). Reports on a method of eliciting and observing certain symbolic hallucination phenomena. Reprinted in (ed.) D. Rapaport, *Organization and Pathology of Thought*. 1951. New York: Columbia University Press.
- SINHA, S. N., FRANKS, C. M. & BROADHURST, P. L. (1958). The effect of a stimulant and a depressant drug on a measure of reactive inhibition, I_R . *J. Exp. Psychol.* **56**, no. 4, 349-54.
- SPENCE, K. W. (1956). *Behaviour Theory and Conditioning*. London: Oxford University Press.
- STARR, K. H. (1958). An experimental study of reactive inhibition and its relations to certain personality traits. Unpublished Ph.D. thesis, University of London.
- STRETCH, R. G. A. (1959). An experimental investigation of curiosity and exploratory behaviour. Paper read at Annual Meeting of the British Association for the Advancement of Science, York, 1959. (In the Press.)

- STRETCH, R. G. A. & KILPATRICK, R. (1959) The effects of certain drugs on maze-exploration and perceptual curiosity in the rat. Paper read to the Northern Branch of the British Psychological Society. (To be published.)
- THURSTONE, L. L. (1924). *The Nature of Intelligence*. London: Kegan Paul.
- WALKER, E. L. (1956). The course and duration of the reaction decrement and the influence of reward. *J. Comp. Physiol. Psychol.* **49**, 167-76.
- WALKER, E. L. (1958). Action decrement and its relation to learning. *Psychol. Rev.* **65**, 129-42.
- WALKER, E. L., DEMBER, W. N., EARL, R. W. & KAROLY, A. J. (1955*a*) Choice alternation. II. Exposure to stimulus or stimulus and place without choice. *J. Comp. Physiol. Psychol.* **48**, 24-8.
- WALKER, E. L., DEMBER, W. N., EARL, R. W. & KAROLY, A. J. (1955*b*). Choice alternation. III. Response intensity versus response discriminability. *J. Comp. Physiol. Psychol.* **48**, 80-5.
- WALKER, E. L., DEMBER, W. N., EARL, R. W. & KAROLY, A. J. (1955*c*). Choice alternation. I. Stimulus versus place versus response. *J. Comp. Physiol. Psychol.* **48**, 19-23.
- WHITE, R. W. (1959). Motivation reconsidered: the concept of competence. *Psychol. Rev.* **66**, 297-333.
- WOLPE, J. (1958). *Psychotherapy by Reciprocal Inhibition*. Stanford University Press.
- ZANGWILL, O. L. (1950). *An Introduction to Modern Psychology*. London: Methuen.

(Manuscript received 6 July 1959)

EMERGENCE OF SET ON THE BASIS OF IMAGINAL SITUATIONS

By R. NATADZE*

State University of Tbilisi, Georgia, U.S.S.R.

The main results are reported of an experimental study in which the writer set out to ascertain the possibility of inducing a set (*Einstellung*) on the basis of some pertinent conceived (and not perceived) situation. In a series of experiments the subject is to imagine, repeatedly and in close succession, that he is lifting a heavy object with one hand, and a light object with the other; or, that he is grasping a large ball with one hand, and a small ball with the other. A fixated set resulting from this imaginal handling of objects differing in size should manifest itself in the subject's subsequent *perception* of objectively equal objects. And so it does: for example, objectively equal weights or equal balls are perceived as unequal. It is concluded that the problem of the feasibility of evolving a set on the basis of imagination has been solved.

I. INTRODUCTION

Man's ability to act not only in conformity with an actual situation but an imaginary one as well is one of the basic peculiarities which distinguish a human being from an animal. Wish-directed activity is 'prospective' in its nature, which necessarily presupposes an ability to act in accordance with a preconceived situation, i.e. a situation that has not yet materialized. Creative activity, particularly the actor's art, presupposes a highly developed ability to act in accordance with an imaginary scenic situation. But since all activity of a living being necessarily presupposes the emergence in the integrated individual of a corresponding state of inner preparedness, i.e. the emergence in the subject of a corresponding set (*Einstellung*), it is to be assumed that one of man's specific properties is his ability to evoke within himself an attitude based on an imagined situation.

The primary aim of the present study was the experimental establishment of the fact of the formation and fixation of a set on the basis of imagination.

II. EXPERIMENTS

In order to be able to make precise comparisons between the manifestations of a set stimulated by an imagined situation and those of a set stimulated by an actually perceived situation (which latter has, for the past thirty years and more, been studied by Uznadze (1939) and his associates in the context of fixated set), we tried so to design the experiments as to make the imagined situation responsible for the emergence of the fixated set reproduce exactly that situation which, in the experiments of Uznadze and his associates, was actually perceived. As to the 'critical' tests, i.e. those in which the fixated set just emerged was to reveal itself, these, as we shall presently see, exactly reproduced the well-known method of Uznadze.

(1) In the first series of experiments the subject, who has his eyes closed, is told to picture to himself as vividly as possible (i.e. to 'experience' as clearly as possible, to delude himself into the appropriate sensation) that he is grasping two wooden

* Professor of Psychology.

balls with his hands—a large one in one hand and a very small one in the other.* Meanwhile the subject's hands actually are resting on the table. After imaginary haptic comparisons of imagined balls have been repeated 15 times we pass to 'critical' tests: for one moment two balls, objectively equal in size, are placed, for comparison, into the subject's extended hands. These 'critical' exposures are repeated until the subject has achieved an adequate perception of the spheres.

Result. The majority of subjects† in the critical tests develop a contrast illusion of perception—the ball on that side on which in the 'critical' tests the smaller ball was visualized is perceived as being the larger one, and vice versa. This is an indication of the fact that a corresponding fixated set had been developed by these subjects in the course of the imaginary grasping of the large and small balls.

An analogous experiment was carried out also in the visual modality. The subjects had a white screen placed before them and were told to imagine that they saw on it, as distinctly and as clearly as possible, two circles—a large one on one side and a small one on the other—to be compared as to size. This procedure having been repeated 15 times in succession, the subject is put to the 'critical' test: two equal circles actually appear on the screen, and the subject has to compare them as to size.

Result. In most subjects a contrast illusion is produced: of the two objectively equal circles, the one on the side on which, in the 'critical' tests, the smaller circle was perceived is now perceived as being larger.

Occasionally the illusion is of an assimilative nature, i.e. that circle is perceived as the larger one which is on the side on which the larger circle was presented in the set-fixation procedure.

The two experiments of this first series have thus established the following fact. As a result of visualizing fifteen times successively a given situation (imaginal comparison of two objects), the subjects evolved a fixated set which so acted upon them that they perceived equal 'critical' objects in an illusive (contrastive for the most part, but sometimes assimilative) way, conformably to the fixated set.

(2) The formation of a set on the basis of a corresponding imaginal situation is therefore an established fact.

The question arises whether a set could not be also evolved on the basis of an imaginal situation, and concurrently with an actual perception of a contradictory content.

To ascertain this a series of experiments has been conducted by us.

A. The subject is told to look fixedly at two wooden 'weights' of equal size (balls with handles at the top) placed in front of him, and at the same time to picture to himself, as distinctly as possible, that one of these weights is very heavy (filled with lead), and the other very light (hollow), and that he is lifting them simultaneously by their handles—one (the very heavy one) with one hand, and the other (the very light one) with the other hand. It is stressed in the instructions that the subject

* The wording of the instruction is of paramount importance. When, in the directions given, the subject is merely required to picture to himself the respective 'objects'—i.e. when he is not *especially* instructed to 'experience' the situation as distinctly as though the two balls were actually being held—the result is often negative: no illusion occurs in the 'critical' tests (see below).

† Subjects in all the experiments reported were adults, ranging from student age to over 50. No differences related to age were noted.

should try 'to feel' with all possible distinctness the difference in weight of the loads that he imagines himself to be lifting.

After 15 successive rehearsals of this imaginal procedure, the subject is given the 'critical' objects: he is to lift (by their handles) two wooden weights, equally heavy objectively, and to compare them as to weight.

It was ascertained that the visual perception of equal objects hinders the subjects in the extreme in conceiving these objects as being different and, besides, sharply distinct in weight, the difficulty making itself yet more felt while the subjects are imagining that they are lifting these different weights. Many subjects try to turn their heads away or close their eyes.

As a result of an experiment made with 60 subjects it has been found that 41 subjects (i.e. 68 %) showed a set-induced illusion, i.e. loads of objectively equal weight were perceived as being different in weight.

Thus a fixated set was evolved in these experiments on the basis of a kinesthetic image of objects differing in size—an image that had been contradicted by the visual perception of the equality of these self-same objects.

B. An experiment in which an illusion of kinesthetic perception, i.e. an illusion of weight perception, is evoked on the basis of the visual perception of magnitudes has long been known to psychological science. What we have in mind is the phenomenon widely known as the Charpentier illusion.

In this experiment, based as we know on the visual perception of a large and of a small cylinder, their objectively equal weight is, in lifting them by strings attached to them, perceived as unequal, the larger cylinder seeming to be lighter.

This should lead up to the following question: will the set-induced effect of the visual perception of the Charpentier cylinders be disturbed by the effects of a preliminary visualization of their weight as being opposed to their volume, this latter being perceived visually?

The experiment. The subject is instructed to keep his eyes on the Charpentier cylinders while trying at the same time to picture to himself that the smaller cylinder is much heavier than the larger one, that the smaller cylinder is filled with lead, while the larger one is hollow inside. The subject is to visualize as distinctly as possible that he is lifting the cylinders by the strings attached to them—the small but very heavy cylinder with one hand, and the large hollow cylinder with the other.

The subject is not allowed to touch the cylinders, his folded hands resting motionless in front of him. After fifteen rehearsals the subject is put to the 'critical' test: he is given the string ends with the instruction actually to lift the cylinders and compare them as to weight.

Thus, in this experiment the set-induced effect of the visual perception of the difference in size of the Charpentier cylinders is set off against the set-induced effect of the visualization of weight as being opposed to the relationship between the sizes of these objects.

Results. In 61.5 % of our subjects the Charpentier illusion proved to be disturbed, i.e. in the 'critical' tests the illusion of perception occurred on the basis of the set-induced effect not of the visual perception of the sizes of the Charpentier cylinders but of the kinesthetic image of their weight as opposed to this visual perception. The Charpentier illusion was disturbed on the basis of a set stimulated by the visuali-

zation of a corresponding relationship of weights. By the majority of the subjects the larger cylinder was perceived as being heavier; this larger cylinder, in compliance with the instructions, was, in the set-inducing trials, visualized as being lighter, a contrast illusion thus being produced whereby to one part of the subjects the cylinders seemed of equal weight.

As to the assimilative illusion, it is not to be taken into account in this experiment, since this effect and that of Charpentier overlap.

C. The question arises whether a fixated set can be evoked on the basis of an image of such content as would be contradicted by an actual perception in the same modality.

The subject simultaneously lifts two equal wooden balls ('weights' with handles) of identical weight. After he has made sure of the equality of the weights of the loads he has been lifting, he is told to lift the same objects simultaneously but in such a manner as to picture to himself, with the utmost vividness, that the load he is lifting with one hand is much heavier than the load being lifted with the other hand, as if the first ball were filled with lead, and the other hollow. After the subject has succeeded in performing this, he is told to go through 12-15 successive repetitions of the same procedure.

We then turn to the 'critical' test: the subject is, several times in succession, given two cylinders of precisely the same weight, with instructions to compare them as to weight by simultaneously lifting them, for the duration of an instant, by means of strings attached to them.

Of the results of the experiment, it is the basic one that matters in the present context. In the 'critical' test, out of 59 subjects 42 showed an illusory perception of the weight of the cylinders—one seemed heavier than the other. Thus in 71 % of the subjects a fixated set was formed on the basis of mental images even though the imagined content was contradicted by a simultaneously presented real content of the same modality (the perception of this content is blurred in the consciousness by a bright mental image of a different content).

It should be noted that most of the subjects in these experiments achieve the required vividness of image—vividness which has no counterpart in the actually given content—only at the price of great mental exertion, and only after a number of attempts have failed.

D. In the experiment just described a set for inequality of weights stimulated by a corresponding mental image suppresses the set for equality. There is experimental evidence that a set for inequality is induced with greater difficulty and is less consistent than a set for equality. And this naturally poses the question of the possibility of overcoming the set-induced effect of an actual perception of different objects on the basis of visualizing a contrastive relation between these two objects; a more particular question here also comes up: is it feasible, while actually perceiving objects inversely related as to weight, to evoke a fixated set on the basis of mentally picturing a process of lifting objects of unequal weight?

The experiment. In contradistinction to the foregoing the subjects are given two cylinders obviously differing in weight but of identical volume, one of the cylinders being 41 % heavier than the other (132 g. and 186 g.). This difference in weight is so considerable that while simultaneously lifting the two cylinders by their strings all the subjects correctly report which of the two cylinders is the heavier one.

After the subjects have convinced themselves of the said difference in weight they are told to lift the cylinders by the strings so as to visualize a reverse relationship of the loads; i.e. the subject must, with all possible distinctness, picture to himself that the objectively heavier cylinder is unusually light (hollow), and the objectively lighter cylinder is very heavy (filled with lead). When the subject has succeeded in carrying out these instructions, i.e. when he has, in the process of lifting the cylinders, been able clearly to visualize the procedure of load-lifting as set forth in the instructions, he is told to repeat this procedure 12-15 times in succession. Then in the 'critical' test, the subject is given two wooden weights of equal volume and weight, and instructed to compare them as to heaviness by swiftly lifting them by their handles.

General result. A set-induced illusion in this experiment appeared in 40 subjects out of 59, i.e. in 68 % of cases. Thus 68 % of the subjects evolved a fixated set on the basis of the mental image of lifting loads whose weight relationship was the opposite of that of the two objects actually lifted. Understandably, there is no longer any factual perception of the real relationship of loads, for it is eclipsed, so to speak, by the contrary image, after the subject has achieved a sufficiently vivid visualization of the reverse relationship.

It must be noted that the required visualization of load-lifting is achieved by the subjects by dint of tremendous mental exertion and only after 14-25 attempts.

Note is also to be taken of the high rate of assimilative illusions of perception in the 'critical' test, i.e. of the cases where the object is perceived as being heavier on that side on which in the set-inducing experiments the heavier object was imagined to have been perceived. On the average, assimilative illusions (in subjects of various categories) constitute one half of the total of cases of set-induced illusion.

E. In the course of the last two experiments it was found that many subjects, in spite of the instructions they had received to the contrary, resorted to 'acting', i.e. while lifting the objects, they deliberately adjusted their muscle tensions to suit the imagined weights, thus creating an actual kinesthetic impression which facilitated a distinct visualization of loads differing in weight from those actually handled.

With a view to eliminating this element of 'acting' or at any rate to making it easily controllable we tried to evoke a fixated set on the basis of a haptic image of different volumes concomitantly with the haptic perception of equal volumes.

The experiment. The subject is simultaneously given two wooden balls of equal volume and weight and told to make a comparison; he simultaneously grasps them (a ball in each hand) and compares their volumes. After he has convinced himself of their equality, he is instructed to grasp the two balls firmly with his hands and try to imagine as vividly as possible (as if it were a sensation) that one of the balls is much larger than the other; and that one of the spheres is so large that it can hardly be encompassed with the fingers while the other is so small that it is almost lost in the hand. The subject is positively forbidden to move his fingers, to open his hand wider than the ball, etc.; each hand should firmly grasp its respective ball. Only in the case where the subject, despite repeated attempts, is unable to visualize what is required, is he allowed from time to time to move his fingers momentarily—to relax the fingers of that hand in which, according to instructions, the larger ball is to be visualized, and then to retighten his grip.

After 12–15 successive signals from the subject to the effect that he has succeeded in visualizing himself as holding in his hands balls of requisite size, we turn to the 'critical' test: the blindfolded subject is presented, several times in quick succession, with equal balls for momentary comparisons (the experimenter holds the balls by the handles).

General result. In these tests a set-induced illusion corresponding to the visualized relationship of the volumes of the balls appeared in 30 out of 58 subjects. Thus only 52% of the subjects evolved a fixated set on the basis of a haptic visualization of different volumes in the course of an actual haptic perception of equal volumes. It must be noted that in this experiment, too, assimilative illusions are of relatively frequent occurrence.

The fact has thus been established of the formation of a fixated set in the haptic sphere on the basis of a mental image whose content is contradicted by real content of an opposite nature.

It should be noted that by the difficulty of evoking a corresponding image while holding equal balls this experiment by far surpassed all preceding experiments. A considerable proportion of the subjects failed to do the task, i.e. to visualize clearly that they had in their hands balls of different size while actually holding equal ones;* a fixated set corresponding to the image was of course not evolved in these subjects. They were unable to visualize different spheres in their hands so vividly as to allow the true perception of different spheres to be eclipsed by the visualized perception.

(3) To what extent does the imagination-stimulated fixated set affect perception? To what extent does the perception of a 'critical' object change under the effect of a fixated set evolved on the basis of a corresponding image?

The experiment. The subject is instructed to keep his eyes on the two equal wooden cylinders in front of him and to imagine as clearly as possible that he is simultaneously lifting the two cylinders by the strings—the heavy cylinder II ('filled with lead') with one hand, and the light (hollow) cylinder I with the other.

After the subject has performed the task 15 times in succession he is told actually to lift, for a moment, both cylinders by their strings in order to compare them as to weight. In reality the cylinders differ in weight—132 g. and 186 g. Thus, the weight of the heavier cylinder exceeds (by 41%) that of the lighter one. This difference in weight is above the discrimination threshold, so that the subject as a rule easily determines, by lifting, which of the two loads is the lighter one.

The object of the experiment is to check whether the contrastive effect of the fixated set will overcome the sensation of the actual difference in the weight of the cylinders lifted in the 'critical' test. Furthermore, the conditions of the experiment make it possible for us to be concerned with the contrast effect of the set taken by itself, i.e. with the contrastive illusion alone, for the direction of the assimilative effect of the set coincides with the actual difference in weight: under an assimilative action the cylinder that is actually heavier must be perceived as such even though the range of the weight differences felt cannot be measured objectively.

As to the contrastive effect of the set, it may manifest itself in two ways: vigorous action resulting in the perception of a reverse relationship of cylinder weights (i.e. the

* It should be emphasized that this refers to a haptic image of different volumes being grasped in the hands; visually, balls of different volume are readily conceived.

lighter cylinder seeming heavier); or feeble action tending to equate the two actually different weights (the cylinders appearing to be of equal weight.)

Results. Twenty-six subjects out of 58 (i.e. 45 %) displayed a contrast illusion. Six subjects out of these 26 (10 %) perceived the 'critical' objects as being equal, while 20 subjects (34 %) took them to be opposites in weight (contrastively different). In assessing this result it should be taken into consideration that usually in analogous experiments the action of a set fixated on the basis of imagination produces a certain percentage of assimilative illusions which, as was noted above, cannot be taken into account in this last experiment.

Thus illusory perception of obviously different weights under the action of an imagination-stimulated fixated set is an established fact. The force of the action of the imagination-stimulated set reaches such a degree that it overcomes (suppresses) differences of weight that would, under normal conditions, be clearly perceived. An object 41 % heavier than another is perceived as being either lighter (34 % of cases) or of equal weight (10 % of cases).

The percentage of illusory perceptions in this experiment, as was to be expected, proved to be lower than in the experiment with equal 'critical' objects.

III. DISCUSSION

Such are the main results of our experiments. To this it must be added that, according to our results, the ability to evoke in oneself a set on the basis of an image displays great interindividual differences; not all the subjects in our experiments carry out the experimenter's instructions from the very beginning.

Among our subjects, actors skilled in stage impersonation and the more gifted of the students of the Theatrical Institute are in a class by themselves. These subjects, almost without exception, carry out our instructions at once and evolve, with remarkable ease, a set on the basis of an image.

A specially striking difference between actors and those unconnected with the theatre was revealed in the last of the above experiments, in which the subjects were, in the 'critical' tests, presented with objects not of equal but of distinctly different weights. In these tests only 9.5 % of the subjects unconnected with the theatre showed a set-induced illusion in the 'critical' tests, one half of them perceiving the 'critical' objects as being equal, i.e. these subjects responded with a faint illusion, whereas of the subjects skilled in impersonation 80 % showed a set-induced illusion; and what is more, an overwhelming majority of these subjects perceived the objects contrastively, i.e. they perceived the objectively lighter cylinder as being the heavier one, and only by an insignificant minority were these objects perceived as being of equal weight.

We believe that this correlation between facility in evolving a set based on imagination and aptitude for dramatic impersonation lends considerable support to our assumption that dramatic impersonation calls, first and foremost, for an ability on the part of the actor to develop a set corresponding to the visualized dramatic situation. If such a set has originated, the actor's behaviour on the stage proceeds with a naturalness that has the power to convince the spectator.

As to those subjects who failed to carry out our instructions (i.e. failed to visualize

a content not given in perception with such vividness and distinctness as to provide a basis for the emergence of a corresponding imagination-stimulated fixated set), it must be emphasized that in our experiments all the subjects succeeded in developing a corresponding set after continued practice and by carrying out our instructions. The majority achieved this after a week's practice (one session a day); with some it took 2 weeks to do so.

IV. CONCLUSIONS

According to the data obtained in the above experiments the fixated set evolved on the basis of an imagined situation is characterized by a number of peculiarities which distinguish it from the manifestations of a fixated set evolved on the basis of perception. We shall here cite only some of them.

(1) The set-induced action of the imagination is conditioned by a number of factors. The image itself, however vivid and clear, does not evoke a set if the subject has no specific relation to the thing visualized.

(2) There are vast individual differences in the ability to develop in oneself a set based on imagination.

By persons capable of dramatic impersonation, and generally, it would appear, by people pertaining to the type in whose inner life imagination plays a major role, the required active attitude toward the visualized content is developed with particular facility.

The ability to develop in oneself a set based on imagination is a manifestation of the personality type, of the peculiarities of the personality and not of the peculiarities of the image itself (its vividness or the like).

(3) A fixated set developed on the basis of imagination as a rule manifests itself in a relatively fainter manner, and is of a considerably less stable nature than are sets developed on the basis of perception. Owing to a number of factors, the stability of this imagination-stimulated set is subject to great variations.

In a few subjects who showed a consistent illusion the effect of an imagination-stimulated fixated set appeared to be more consistent than the action of a set developed on the basis of perception in ordinary set-inducing experiments. These subjects artificially exaggerated the difference of the objects visualized in the set-inducing experiments and artificially stimulated in themselves the experiencing of the significance of this difference.

(4) In a deliberately developed set, especially when the visualized content is contradicted by what is perceived, it is not infrequent that only a hardly perceptible set-induced illusion is observed in the initial perceptions of 'critical' objects—an illusion which gradually increases in the subsequent perceptions up to the moment when the process of extinction of the illusion sets in.

(5) In developing a fixated set on the basis of imagination its effect proves to be incomparably more irradiated than in developing a fixated set on the basis of actual perception. In our experiments, irradiation from a visual-image set to haptic perception was observed in all, or nearly all (80 %), the subjects in whom a set based on imagination was at all developed: when in the set-inducing experiments the subject visualized balls of different size while in the 'critical' tests he grasped equal balls with

his hands and compared them by touch as to size, transposition of the illusion was observed in 80-100 % of subjects.

(6) In a deliberately developed fixated set having imagination for basis, especially under conditions of such content-visualization as would be contradicted by the actually perceived content, assimilative set-induced illusions are more widespread in 'critical' tests than in evolving a perception-set. This phenomenon is apparently to be accounted for by the fact that the subject continues, 'by inertia', to transform the perception in the same direction as in the set-inducing experiments when, with great effort, he overcomes the adequate perception of real objects and visualizes them as being greatly altered.

(7) The ability to develop in oneself a fixated set through imagination is highly subject to practice effect: as a result of practice with the tasks set in our experiments all our subjects succeeded in developing a fixated set on the basis of imagination.

The above series of experiments, conducted in the various modality spheres, have thus demonstrated the feasibility of evolving a set on the basis of a mental image whose content is contradicted by actual perception.

What are the factors giving rise to this phenomenon? What are the factors underlying the ease with which an image-based set is developed in some subjects? A special experimental investigation has been devoted to this problem by the present writer (Natadze, 1958). Space forbids dwelling on it now.

REFERENCES

- NATADZE, R. (1958). On some factors in the formation of attitude based on an imaginary situation. *Voprosy Psikhologii*, no. 3, 35-46.
UZNADZE, D. (1939). Untersuchungen zur Psychologie der Einstellung. *Acta Psychologica*, 4, 323-60.

(Manuscript received 17 November 1959)

THE STROBOSCOPIC PATTERNS

III. FURTHER EXPERIMENTS AND DISCUSSION

BY J. R. SMYTHIES*

The Psychological Laboratory, Cambridge

This paper completes the report on the stroboscopic patterns. In two previous papers in this *Journal* (Smythies, 1959*a*, *b*) an account was given of the phenomenology of the dark and bright phases of the stroboscopic patterns—the subjective sensory patterns evoked by illuminating the retina with an intermittent stimulus. This paper reports some further experiments that explore the parameters that affect these patterns and a discussion is presented of the various hypotheses that have been advanced as to their origin.

I. FURTHER EXPERIMENTS

(1) *The effect of stimulus form.* It was noted in part II of this paper that the patterns evoked in each subject by an electroencephalogram electronic photic stimulator and those evoked by a rotating sectored disk in a beam of light were very similar (except in the case of only one subject (no. 6) out of thirty-five). These two modes of stimulation have quite different wave forms—very short flashes and a light-dark ratio of 1:3 respectively. A third wave-form tried was a sine wave produced by rotating one polaroid filter in front of another stationary one. This produced essentially the same type of patterns as the two former types of stimulation in four subjects. Furthermore, if an Archimedes spiral is rotated rather quickly and then 'observed' with relaxed accommodation from 3 to 4 in. away, or if a large fan is rotated at the right speed and observed against the sky, the patterns usual for the subject can be evoked. This indicates that the particular form of the stimulus is not important so long as there is a vigorous temporal alternation of light and dark of some kind and this need not be uniform all over the retina.

(2) *The effect of different lenses.* The effect of interposing strong convex, concave and cylindrical lenses between the eye and the stimulus source was determined. None of these made any difference to the form of the patterns; a result to be expected from the featureless stimulus used.

(3) *The effects of in-phase and out-of-phase stimulation of the two eyes.* Using a rotating disk and a series of mirrors, two beams of light 90 degrees out of phase in their intermittence were cast on an opal screen behind which the subject was sitting. These two beams produced contiguous squares of light on the screen wider than the distance between the subject's eyes. Thus merely by moving his head he was able to compare in-phase and out-of-phase stimulation of his eyes. Thirteen subjects were asked to report whether these two modes of stimulation were the same or different. They all reported that they were different. Eight reported a more elaborate pattern and three a different pattern with the out-of-phase stimulus, and two did not specify the difference. The experimenter noted that the out-of-phase pattern resembled the in-phase pattern of about twice the frequency of stimulation.

* The experiments reported in this paper were conducted during the tenure of a Nuffield Fellowship in medicine, and the work was completed at the Galesburg State Research Hospital, Galesburg, Illinois.

(4) *The effect of suddenly doubling or halving the flash frequency.* In this experiment two slide projector lamps were shone through opposite sides of the sectorized disk on to the same area of the opal screen. By switching one of these lamps on or off the frequency of the flash could be suddenly doubled or halved. Twelve subjects were used. The commonest report was that the same pattern remained, but that its elements became fewer and coarser on halving the frequency and more numerous and finer on doubling it. This applied to the following patterns: mosaic pattern of squares, vertical lines, checkerboard, cross, concentric circles and honeycomb (for examples see Figs. 1-6). Once the texture of the field was merely reported to be 'thinner, finer and more tenuous at *F* (fast frequency)' and 'thicker and grosser at *S* (slow frequency)'. At other times the pattern changed altogether (e.g. grid to whirlpool) or merely vanished. There was no effect on the dark-phase patterns which were the same at all frequencies tried. Once in one subject thin vertical lines at *F* changed to thin radial lines at *S* and on another occasion this effect was partly obtained. In the case of another subject the lines actually got thinner at *F* than at *S* when the subject was fatigued from long observation, although this subject, when not so fatigued, showed a great preponderance of thick crosses at *S* and thinner ones at *F*. In one subject the diamond spiral consistently had a thicker band and fewer turns at *F* than at *S*. Table 1 summarizes these results quantitatively.

Table 1. *The effect of suddenly doubling or halving the flash frequency*

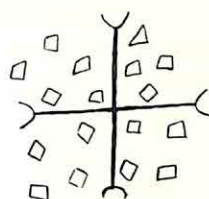
Result of change of frequency	Number of reports
The main result described	66
The reverse effect (fine at <i>S</i> , coarse at <i>F</i>)	13 (6 with fatigue; 7 the diamond spiral)
Pattern went	23
Pattern changes in unspecified way	35
Same pattern remains unchanged	5
Some element of main sequence	4
Pattern changes in some other way	3 (grid to squares)
Total	149

(5) *The effect of random flicker.* A random flicker was obtained by driving the photic stimulator from a Geiger counter responding to a radioactive source. By varying the distance from the detector of the counter to the source the average rate of the random flicker could be raised or lowered.

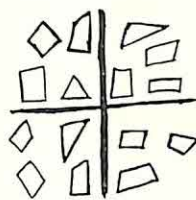
(a) *Subjective observations.* The patterns appeared to be trying to follow any regularity there was in the random flicker. In such a stimulus there will be short runs of flashes close together and some periods of more infrequent flashes. During these periods the patterns took on the typical aspect of the patterns associated with *F* and with *S* stimulation respectively. The change from one to the other was very rapid and in between there were periods of rapid change of patterns in a more confused manner. No new appearances were observed.

(b) *Objective studies.* An attempt was made to set up a matching experiment as follows. The subject was given a standard regular flicker with which he had to match the random flicker by observing consecutive runs of each. The experimental arrangements were as described for the strobe lamp (photic stimulator) in part I of this

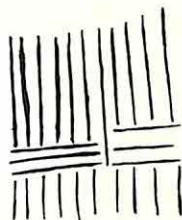
paper. The subject altered the rate of the random flicker by moving the radioactive source nearer and farther from the detector of the counter and he could change from one mode of stimulation to the other merely by throwing a switch. Eight speeds from 6 to 20 f./s. were set by the experimenter in steps of 2 as the standards. Then the controls were reversed and eight different average random frequencies were set (by the experimenter) as standard and the subject matched these by selecting the appropriate regular matching frequency by turning the dial on the strobe lamp. The



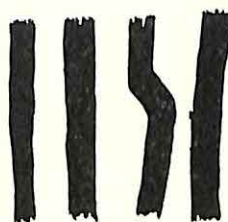
1



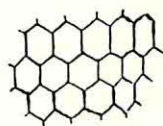
2



3



4



5



6

Figs. 1-6. Three examples of the effect of suddenly halving or doubling the flash frequency. The odd numbers represent the higher frequency patterns and the even numbers lower frequency patterns. Figs. 1, 2 are from subject 16; Figs. 3, 4 from subject 34; and Figs. 5, 6 from a subject outside the series.

matching was done by adjusting the frequency of the unknown to give the stroboscopic patterns most like those observed with the standard. Thirty-two such runs were done with seven subjects. As a control the subjects repeated these same experiments, but were told to try and match on a basis of the frequency of the flicker and to ignore the form of the patterns. In the former experiment the subjects were sitting with their faces right up against the opal screen (where the patterns are most obvious) and in the latter experiment they were sitting about 3 ft. back from the screen where the patterns are less clearly visible but the rate of flicker is just as noticeable.

Fig. 7 shows the result. The differential rate of increase of matching regular and random stimuli is to be expected as the higher the average rate of a random flicker the more small groups of flashes there will be closer together than the c.f.f., and thus which will appear as a single flash. The fact that the matching on speed is just as good as the matching on pattern means, however, that this experiment fails to give any objective indication of the relation between the type and form of the patterns obtained with regular stimulation and those obtained with random stimulation, because one cannot be certain that the clues of frequency of flash were not used even in 'pattern matching'. Some subjects felt subjectively certain that they really were matching on the form of the patterns but others found the experiment very difficult. The most that can be said is that the patterns obtained with random stimulation resemble those obtained with regular stimulation in certain ways.

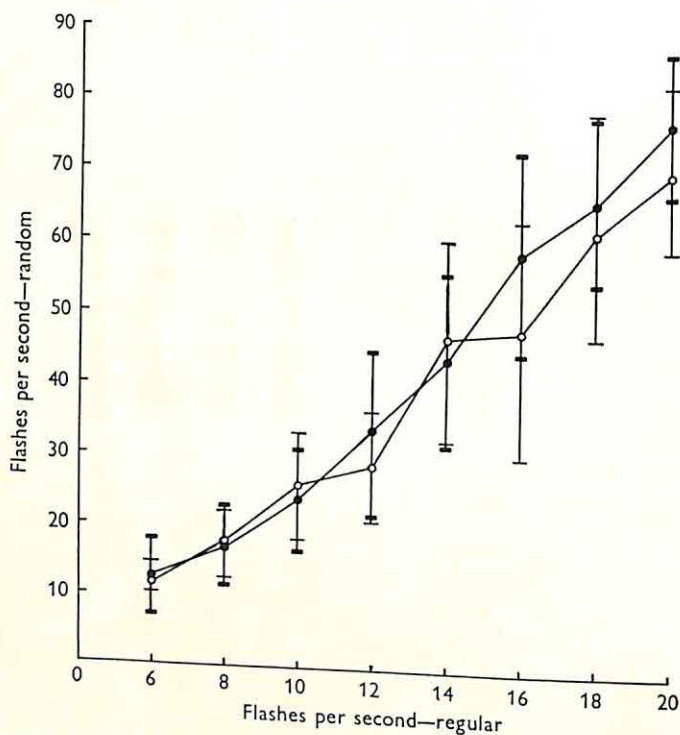


Fig. 7. The matching levels for comparing random and regular flicker. Average deviations are given.

II. DISCUSSION OF EXPLANATORY HYPOTHESES

A number of hypotheses have been put forward to account for the stroboscopic patterns. The problem that we have to solve is why, when we look at a uniform field under intermittent illumination, we do not merely see the stimulus but see instead these complex and interesting patterns.

(1) *The first hypothesis.* The explanation usually put forward by early workers is that these patterns represent entoptically visualized retinal structures. Any radial system of lines could be ascribed to the retinal blood vessels; hexagonal patterns to the pigment cells of the retina; central diamonds to the macula pigment, etc. However, these patterns are mainly too regular and geometrical to be biological structures,

and moreover they may be observed to rotate in a regular fashion and to execute other gross movements, which precludes their being retinal structures. The only pattern met with in this investigation that merited such an explanation was the central blob or disk. This was particularly prominent in blue light and might have been the entoptic visualization of the yellow macular pigment. One could not be certain about this as, on occasion, a bright yellow central disk could be obtained in yellow light. Nevertheless, such an origin for this pattern could not be eliminated and thus all central blobs or disks were excluded from the pattern counts.

(2) *The second hypothesis.* This was first put forward at some length by Grey Walter (1950). He represents the patterns as interference phenomena produced in a scanning mechanism attempting to deal with an intermittent signal and equates them with the patterns that can be produced on the screen of a television set by illuminating the studio with an intermittent light. He makes the significant point that the *form* of any such patterns will be determined by the particular form of the scanning raster (the arrangement of scanning beams used to build up the final picture) employed. Now clearly, if this hypothesis is true, the physiologically important information of the form of the actual raster employed by the physiological 'scanning' mechanisms can be obtained by the study of these patterns. For example, a linear scan might be expected to give rise to patterns of segments of straight lines, horizontally orientated in the case of a horizontal scan or vertically orientated in the case of a vertical scan. Similarly, a polar scan would give rise to patterns compounded of segments of curved lines, and a radial scan (of the type commonly used in radar) would give rise to patterns compounded of radially arranged lines. One observation and two series of my experiments are relevant to this hypothesis. The observation is that the types of pattern found do seem to fall into the three classes suggested above from a consideration of the commonly used rasters—i.e. patterns of classes 3, 4 and 6 lend themselves readily to a derivation from a linear, radial and polar scan respectively. Class 5 patterns could possibly arise from a linear scan also as, for example, herring-bone patterns are a common aberration seen on ordinary television screens. This correlation may, of course, merely be a coincidence but it does suggest that a study of such interference patterns obtainable with the various types of scan might be of interest to a communications engineer in this context.

The experiment described in section (4) of this paper gives the result that would be predicted by a scanning hypothesis since the grain of an interference pattern produced by intermittent excitation of a scanning mechanism will be determined by the frequency of the intermittent stimulation in the manner found. The other experiment has been described which compares the patterns produced by a random flicker with those produced by a regular flicker. The similarity of these two types of pattern, particularly in the case of the vertical line patterns reported by subject 34 (one of our most skilled observers) would be somewhat unexpected under any simple scanning hypothesis. The scanning hypothesis has been under adverse criticism recently (Mackay, 1953; Walsh, 1952). However, it should be noted that these criticisms and the experiments on which they are based refer to the particular hypothesis that the alpha-rhythm of the E.E.G. is a necessary feature or product of such a mechanism. The possibility remains that a scanning mechanism might operate in the visual

system quite unrelated to the alpha-rhythm. So it may be concluded that this hypothesis is still a valid one to account for the stroboscopic patterns, but much more work must be done before the matter can be decided.

(3) *The third hypothesis.* This was put forward by H. Barlow (1956, personal communication) as follows: 'the flickering field evokes very vigorous activity in the visual pathways, but the stimulus is almost totally unfamiliar to the subject. In communicating his sensations the subject has to refer to causes of his sensations which are ordinary and familiar, since these are what his vocabulary is adapted to. The prevalence of moving lines and contours in the subject's descriptions might be explained along these lines, since these are the commonly occurring, simple stimuli which will cause vigorous and sustained activity in the "on" and "off" units of the retina, but it must be admitted that there are other, specific features of the descriptions which are not accounted for.' Thus the functional 'on' and 'off' elements of the retina are normally stimulated by contours of darkness and light—by pattern—crossing the retina during the ceaseless movements of the eyeball and environment. Thus it may be supposed that the brain interprets activity in the fibres coming from these retinal elements in terms of pattern. The intermittent means of illumination here employed may be supposed to generate a good deal of fairly regular activity in the on-off cells, fibres and their associated analysing mechanisms. Thus the aberrant form of the perceptions induced by this mode of stimulation might give us information about the manner of functioning of these mechanisms. Normally the retinal mechanisms code, and to some extent analyse, the information received at the retinal surface. This coded information is transmitted to the cortex and is there decoded and analysed to yield the final perception of the object. Thus the aberrant perceptions represented by the stroboscopic patterns might arise from faulty coding, transmission or analysis (or some combination of these) of information presented in this case in a form with which the perceptual mechanisms are not designed to deal. A regular grid, a whirlpool, a family of parabolas, arising out of a simple stimulus of a spatially uniform field under conditions of temporally intermittent illumination, may reflect some fundamental specification of the visual mechanism. A similar hypothesis has been put forward by Walsh (1952). One way of thinking about this might be as follows. The pattern analysing mechanisms of the brain are here presented with information—with coded signals—of a totally unfamiliar kind. They receive a powerful intimation of *pattern* but without any particular pattern being specified. The brain then makes 'hypotheses' and the ever-changing play of patterns may represent the various 'hypotheses' tried and discarded in turn. Against this, however, is the fact that the mechanisms never seem to learn their error. The unsatisfactory 'hypotheses' continue to be produced in spite of the conscious knowledge of the subject that he is really seeing only a flickering light. The analysis may, however, be conducted at levels of the nervous system not amenable to conscious control. But this hypothesis does not explain why some people's patterns should be so stable and unchanging, unless we suppose that the selection of pattern to fit the stimulus is not entirely random but is modified to some extent by the functional characteristics of the analysing mechanisms to give a bias, to a greater or lesser extent in different people, to certain kinds of pattern. It may then be possible to deduce these functional characteristics from the form and behaviour of the patterns. My task here is to

present the latter facts and to sketch possible means of bridging this gap. But the particular deductions can only be made by someone familiar with the detailed operation of telecommunication mechanisms. However, it is possible at this stage to draw attention to what may be important clues for this. These are: (i) The fact that the patterns are not compound of straight *and* curved lines as connected elements. This links up to some degree with one feature of one of Deutsch's (1955) models of cortical mechanisms, which he states would experience some difficulty in distinguishing straight and curved lines. This feature of the stroboscopic patterns suggests that the brain handles straight and curved lines differently. (ii) The effect on the patterns of dark and light adaptation. (iii) The effect of suddenly doubling and halving the flash frequency. (iv) The difference between the initial and the developed patterns. (v) The fact that the parallel vertical line pattern is obtained only with unioocular stimulation and the other constant features of the phenomenon described.

(4) *The fourth hypothesis.* This was suggested by P. E. K. Donaldson (1956, personal communication) in the form that these patterns represent the formation of corresponding domains among retinal and/or cortical neurones along the lines suggested by Cragg & Temperley (1954). Barlow (see above, p. 252) notes that the lateral inhibition in the retina makes it likely that the conditions are right for domain formation, and the differential spread of excitation and inhibition around neurones may lead to the complex pattern formation. One can imagine that each successive wave of excitation from the retina subjected to this intermittent stimulation may interact in the cortex to form complex 'wave' patterns much as the vibrations of a bow drawn across the edge of a tray leads to complex patterns in sand laid on the tray.

(5) We can note also that there are two purely physiological features that may contribute to some features of the patterns. The first is that the neurones of the fovea conduct faster than those of the retinal periphery and clearly this could lead to a flashing light being represented by a series of expanding concentric circles. Secondly, the different organization of the central retina in terms of the relative proportion of rods and cones and the different number of rods and cones connected to each ganglion cell may explain the different organization of the central and peripheral fields that is so often found. The initial patterns (typical of dark adaptation) tend to linger in the periphery of the field, while the developed patterns appear in the centre of the field. This may be correlated with the fact that the periphery of the retina is more adapted to scotopic vision. However, one should note that the developed patterns cannot be accounted for simply by the more light-adapted state of the eye. For, if adaptation is made to a steady light and then the eye exposed to flicker, one does not leap straight into the developed pattern but rather, in many cases, to a mosaic pattern peculiar to this state. This suggests that the prolonged exposure to intermittent light produces patterns of organization of neurones different from those produced by exposure to steady light.

The difference between the patterns following dark adaptation and those following light adaptation together, in the case of the experimenter, with the preponderance of zig-zag line patterns in dim lights and maze patterns in bright lights, suggests a possibility that the I.P. or line patterns are a function of the rods and the D.P. or maze patterns are a function of the cones. In this case the former might be observed predominantly in blue light (stimulating mostly rods) and the latter in red light

(stimulating mostly cones). However, this was not found to be the case and it is more probable that the I.P. and D.P. patterns, and the line and image patterns, reflect different organizational patterns of neurones, rather than that they are the product of the photopic and scotopic mechanisms respectively.

(6) *Correlation with patterns produced by other means.* Two other means of producing hallucinatory patterns have been reported to give rise to similar forms as the stroboscope. The first is by means of moderate and long-continued uniform pressure on the eye. Purkinje (quoted by Helmholtz, 1924) 'described fine quadrangles in regular array, on which there were either stars with eight rays, or dark or bright rhombs with vertical and horizontal diagonals and the patterns were surrounded by alternately bright and dark bands'. Helmholtz (1924) described 'the background of the visual field is usually finely patterned at first, but in the most manifold way and in very different colours. Frequently, it is as if the field were strewn with fine leaves or covered with moss; then presently they look like bright brownish yellow quadrangles everywhere with fine line patterns; and at last they usually develop in the form of dark lines on a brownish yellow background. Sometimes they assume very complex star-shaped figures, and sometimes they are in the form of an inextricable labyrinth or maze, which seems to be waving or flowing continually.' I have tried to repeat this but have only produced flowing masses of colour before the pressure became too uncomfortable. The second means of inducing geometrical hallucinations of this type is by taking mescaline. Klüver (1942) reports the following 'form constants' found in the geometrical patterns typical of the initial stage of mescaline intoxication: (a) *grating, lattice, honeycomb, chessboard, fretwork, filigree*; (b) *cobweb*; (c) *tunnel, funnel, alley, cone, vessel*; (d) *spiral*. (The items in italics are common to the stroboscopic patterns.) Notice the same grouping into linear, radial and polar patterns. Klüver notes that 'some or all of the form constants found in mescaline hallucinations are also found in certain hypnagogic hallucinations, in entoptic phenomena, in the visual phenomena of insulin hypoglycaemia, and in phenomena induced by simply looking at disks with black, white, or coloured sectors rotating at certain speeds'. And again: 'We wish to stress merely one point, namely, that under diverse conditions the visual system responds in terms of a limited number of form constants.' Two further points may be made concerning the relation of the mescaline and the stroboscopic patterns: (i) The mescaline patterns tend to be more regular and more ornate. This tendency increases as the hallucinosis develops, the earlier lattices and checkerboards giving way to elaborate arabesques, filigrees, persian carpet patterns, tapestries, etc., and these in turn to the fully formed scenes typical of the height of the condition. (ii) The effect of the intermittent illumination potentiates the effect of mescaline and hallucinations of the mescaline type may be induced by its use with a dose of mescaline (200 mg.) that would not ordinarily induce hallucinations.

REFERENCES

- Cragg, B. G. & Temperley, H. N. V. (1954). The organisation of neurones: a co-operative analogy. *Electroenceph. Clin. Neurophysiol.* **6**, 85-92.
- Deutsch, J. A. (1955). A theory of shape recognition. *Brit. J. Psychol.* **46**, 30-7.
- Grey-Walter, W. (1950). In *Perspectives in Neuropsychiatry* (D. Richter, ed.). London: Lewis.
- Helmholtz, H. von. (1924). *Physiological Optics*. Rochester: Optical Society of America.
- Klüver, H. (1942). Mechanisms of hallucinations. In *Studies in Personality*. New York: McGraw-Hill.
- Mackay, D. M. (1953). Some experiments on the perception of patterns modulated at the alpha frequency. *Electroenceph. Clin. Neurophysiol.* **5**, 599-62.
- Smythies, J. R. (1959a). The stroboscopic patterns. I. The dark phase. *Brit. J. Psychol.* **50**, 106-16.
- Smythies, J. R. (1959b). The stroboscopic patterns. II. The phenomenology of the bright phase and after images. *Brit. J. Psychol.* **50**, 305-24.
- Walsh, E. G. (1952). Visual reaction time and the alpha-rhythm; an investigation of a scanning hypothesis. *J. Physiol.* **118**, 500-8.

(Manuscript received 28 July 1958)

IMPERCEPTION OF PHI: SOME FURTHER DETERMINANTS

By VERNON HAMILTON

*Birkbeck College, University of London**

An apparent movement situation was presented to control and neurotic groups. Perception of the phi phenomenon was postulated to be associated with neurotic mechanisms and with behaviour operationally defined as conflict and ambiguity avoidance, and was considered to express a generalized perceptual attitude of the individual. Neurotics were found to have lower mean movement intervals than control subjects, and to find it more difficult to perceive the phi phenomenon. The latter tendency was particularly marked in obsessionals and hysterics. Perception of phi was found to be associated with age of subject, becoming more difficult with greater age. Imperception of phi among the neurotic sample remained greater than in the control group after partialling out age. The findings are interpreted as supporting two hypotheses derived from a set of postulates linking a perceptual attitude of ambiguity avoidance with ego defences.

I. INTRODUCTION

By contrast with the majority of studies devoted to considerations of apparent movement phenomena, the present paper will be mainly concerned with factors that appear to operate against the subjective experience of phi. In the phi phenomenon rapid and successive presentation of two stationary stimuli gives rise to an experience of spatial displacement of a single stimulus leading to a perceptual experience of movement. For many years it had been assumed that the movement phenomenon was dependent on an optimum time interval between the members of a pair of stimuli. Subsequent work centred largely on the interaction of spatial, brightness and temporal parameters of the stimulus situation, and the influence of environmental context and subjects' 'set'.

Most studies have specified the conditions under which stroboscopic movement of lines, circles or lights could occur, and the relative ease with which movement neither inappropriate nor incongruous for certain stimulus objects could be simulated. The findings of Klein & Schlesinger (1951*b*) and Jones & Bruner (1954) among others, indicated that stimulus context, expectancy and meaning are relevant variables in the experience of apparent movement, and all studies assumed the 'inevitability' of the movement experience under appropriate conditions.

Explanations of stroboscopic movement appear generally to have been insufficiently tested, and this is particularly true for simple and 'non-pictorial' stimuli. The value of a physiological hypothesis could be considerably enhanced by considering, for example, variations due to changed conditions of neural conductance which may be expected to occur under the influence of drugs. Studies by Gantenbein (1952) and Meili & Tobler (1931) suggested that the phenomenon may be experienced more easily by children, but their work did not extend to a study of older age groups. Klein & Schlesinger (1951*b*) have related consistent individual attitudes towards instability to the perception of apparent movement, and several investigators (Bender & Teuber, 1949; Werner & Thuma, 1942; Shapiro, 1954; Brenner, 1956) have

* Now at Psychology Department, Springfield Hospital, London, S.W. 17.

attempted to explain phi within basically similar frameworks of cortical excitatory-inhibitory states. Apart from the work of Klein no attempt appears to have been made to relate the phenomenon to any general cognitive principles governing perception, and few links appear to have been attempted between what are usually and somewhat vaguely called 'aspects of personality' and the perception of apparent movement.

The experimental findings that will be reported were obtained in a larger investigation which has formed the subject of previous papers (Hamilton, 1957 *a, b*) on some of the correlates of avoidance of perceptual ambiguity. This investigation had set out to study the degree of generality of individual responses to a variety of operationally defined ambiguous situations, and to determine whether there were any differences in this generalized approach between neurotic and normal subjects, and between the neurotic subcategories.

The previous findings can be only briefly summarized. Eleven test situations were presented to subjects, each of them considered to contain operationally defined aspects of perceptual ambiguity. The subjects were a normal control group and groups of hospitalized neurotics: obsessionals, conversion hysterics and patients presenting predominantly anxiety symptoms. The groups were not significantly different in age, intelligence and sex. Three tests were concerned with sensory discrimination. A series of weights had to be compared with a standard weight. Four of the stimuli were clearly unequal to the standard, in order to maintain a 'set' for difference judgments, the remaining twelve weights were made equal to the standard. Subjects were told, however, that none of the stimuli was equal to the standard, but that the difference would sometimes be very small indeed. The stimuli had to be sorted into three categories: 'Heavier', 'Lighter', 'May be Heavier or Lighter, I can't decide which'. Scores in the latter category were taken to indicate degree of avoidance of ambiguity. Procedures were similar for length and brightness discrimination tests with appropriate changes in presentation and response categories.

The battery also contained three series of twelve drawings each. These will be referred to as ambiguous pictures I, II and III in the text below. In Series I some drawings were clearly sheaves of corn, others were equally clearly drawings of trees, other drawings were of indefinite design or contained in the same drawing characteristics of both objects. Series II was similar to Series I except that the two objects in the drawings were a house and a car. Series III differed from Series I and II in that in this case a drawing of a car changed by degrees into a drawing of a dog, but there were also a number of rather vague and indefinite designs. In each series the drawings had to be sorted by subjects into the following groups: 'Object 1', 'Object 2', 'May be either Object 1 or Object 2, I can't decide which'. The final score in the latter category was obtained after subjects had removed from the 'either...or' category, any drawings which they thought really represented 'neither Object 1 nor Object 2'.

A further test consisted of forty-four wooden blocks incorporating a variety of triangular, rectangular and circular features both singly and in combination for different blocks. The objects had to be sorted into 'Rectangular' and 'Circular' categories, into a group for those which subjects considered could be either and they could not decide, and a fourth group for blocks which ought to be grouped into separate categories.

A scale with items printed on separate cards was included to obtain information on subjects' reaction to socially ambiguous situations such as 'A solution to any problem is either right or wrong'. Items were 'posted' into boxes marked 'True' or 'False'.

The battery also contained two reversible perspective drawings, the Necker Cube and the Rubin 'Double Cross', where the number of perspective oscillations was recorded, and low scores were considered to indicate avoidance of perceptual ambiguity. The majority of correlations between the tests were positive and nearly half of them were significant at the 1 or 5 % level of significance. These findings were taken as evidence for a general principle of cognitive control expressing itself in this case in individually consistent approaches to perceptual ambiguity and conflict. Intergroup comparisons indicated that neurotics avoided ambiguous situations more markedly than the control group, and that obsessionals and conversion hysterics tended to behave somewhat similarly in that they avoided the ambiguous situations and response categories to a greater extent than did anxiety neurotics.

The degree of individual consistency in response to indeterminate and unstructured stimuli, suggested the relevance of the classical apparent movement experiment in this kind of investigation. The framework of findings summarized above, taken in conjunction with the concept of intolerance of ambiguity first outlined by Frenkel-Brunswik (1949), and the theoretical formulations concerning generalized principles of cognitive control expressed through consistent perceptual attitudes, put forward by Klein (1951*a, b*, 1953, 1954), suggested five postulates. (1) Traditional apparent movement experiments present to subjects a stimulus situation of considerable perceptual ambiguity. (2) Individual responses in apparent movement experiments would be considerably influenced if not mainly determined by a generalized, customary and preferred method of dealing with all kinds of perceptual ambiguities. (3) Preferred methods of response to perceptual ambiguity, instability and conflict as differences in *the need to avoid* perceptual ambiguity, and emotional ambiguity and conflict would be determined by the need of individual personalities to control internal conflict would be determined by the need of neurotic patients to control internal conflict. (4) The greater need of neurotic patients to control internal conflict would predispose them to avoid the conflictful experience of accepting reality-defying perceptual organization. (5) The avoidance of perceptual conflict and ambiguity would be most important for the most highly 'defended' neurotics—obsessionals and conversion hysterics.

There is, of course, relatively little experimental evidence to support some of the assumptions in the suggested postulates. In this, as in other cases, the postulates are only as good as the degree of support obtained for the hypotheses derived from them. It was proposed to test two particular hypotheses from amongst a number of equally possible ones, not because it was intended to ignore or supplant the evidence from physiologically or neurologically determined hypotheses, but in order to demonstrate the influence of yet another parameter in a complex phenomenon. The following two hypotheses were therefore submitted to experimental test. I. Neurotic subjects and particularly obsessionals and conversion hysterics would find it more difficult than normal control subjects to perceive apparent movement. II. Subjects who rejected the ambiguity of the apparent movement situation by being unable to perceive movement, would also show a generalized tendency of ambiguity avoidance in other unstructured, indeterminate stimulus situations.

II. METHOD

Procedure

The nature of the present investigation required maximal ambiguity of the stimuli. It was, therefore, decided to present the pair of stimuli under conditions of single exposure rather than under the more usual cyclical conditions, and to present 'meaningless' blobs of light rather than objects which are customarily seen in motion. Prior to the experiment subjects were shown that there were in fact two stimuli. It may be considered, therefore, that the total situation was loaded against the perception of apparent movement, and in the direction of greater ambiguity. Since the investigation employed experimentally unsophisticated subjects it was thought to be too difficult for them to distinguish between pure phi—the actual and realistic experience of movement of a single object—and optimal movement, i.e. an experience of succession or simultaneity of stimuli accompanied by a superimposed impression of something moving. (This distinction has been considered important in the past especially by Wertheimer.) Inquiry showed that of those reporting an impression of movement, one-third reported optimal movement, about a third thought they experienced pure phi. The remainder reported a definite impression of movement without being able to decide whether it was pure phi. No movement response was accepted as such unless a subject could also confirm his report of movement by describing with his hand correctly the direction of illusory movement (left to right from his seat) determined by the apparatus. Thus reports of experience of movement in a right to left direction were not counted as experience of phi.

By contrast with most previous experiments, including some recent work by Jeeves & Bruner (1956), the constant method rather than the limiting method of presentation was adopted for determining the movement range and the mean movement interval. While a decisive experiment remains to be done, it may be safely postulated that ascending and descending series of time intervals produce perceptual expectations which are likely to be reduced when interval values are given randomly. A fixed but random series of time intervals separating the appearance of the two stimuli was selected on the basis of a trial run for a small number of control subjects. Each interval was presented five times before changing to the next interval in the series which was then in turn presented five times. The whole series of thirteen stimulus intervals was presented three times in succession. The intervals in millisees were as follows: 153, 412, 0, 102, 357, 205, 51, 306, 127, 255, 168, 76 and 25. Subjects were introduced to the experiment by showing them that there were actually two lights which could come on together or in succession. Movement was defined to them as the impression '... that the spot of light that comes on first moves over to the position occupied by the second one'.

Apparatus

The apparatus used in the investigation consisted of two units. The commutator was an electrically driven cam on which two pairs of 'make' contact springs were riding. While the contact springs moved through the 'dip' of the cam, the respective lamps were illuminated. The stimulus unit was made up of a ground-glass screen measuring 10 × 6 in. behind which two torch-bulb units, *A* and *B* (2.5 V., 0.3 A.) could be independently moved in a horizontal plane. Each of the lights was electrically connected to one of the contact springs on the commutator. The distance between the lights *A* and *B* (*s*) = 6 cm., the exposure time of the stimuli ($t_1 = t_2$) = 0.075 sec., the stimulus diameter = 0.90 cm., the time interval between the first and the second member of the pair of stimuli was variable between 0 and 412 msec., while the time interval between the single exposure of the two lights and the next presentation of the first stimulus was 3 sec. Room illumination was kept within the range of 1½–5 f.c. and the exposure time of the stimuli and the distance between them was kept constant throughout. The apparatus employed was mechanical and of rather primitive though efficient design, and being rather noisy it necessitated a variety of damping devices. Departure from the usual methods of cyclical presentation under darkroom conditions, and the apparatus and method of presentation employed here, were factors likely to yield fairly specific data due to their unorthodoxy. The experimental conditions may be said to err, however, on the side of greater perceptual ambiguity as required by the general framework of the investigation, and they were, of course, homogeneous for all experimental samples.

Subjects

The experimental group consisted of sixty-two symptomatically diagnosed neurotic in-patients from a number of mental hospitals near London. Twenty were obsessionals, twenty were conversion hysterics and twenty-two were patients with predominant symptoms of anxiety. The control group consisted of forty manual, clerical and nursing personnel from the same hospitals and a University College. The groups were adequately equated for intelligence and sex (Hamilton, 1957b). The differences in mean age between controls and neurotics, and between the neurotic subsamples did not approach statistical significance. While it had been hoped to include a small group of patients from a general medical and surgical unit in the experiment, the relevant hospital authorities were not able in this case to provide the necessary facilities.

III. RESULTS

The experimental data will be considered separately for subjects who perceived apparent movement and for whom individual differences in mean movement intervals are available, and for subjects who failed to perceive ϕ . Table 1 shows the different

Table 1. Perception/imperception of ϕ

	Total	Controls	Neurotics	Anxiety states	Obsessionals	Conversion hysterics
Nos. 'seeing'	64	32	32	15	8	9
% 'seeing'	63	80	52	68	45	40
Group N.'s	102	40	62	22	20	20

proportions of subjects perceiving movement. Testing the null hypothesis for differences between the samples with D.F. = 1, the following χ^2 and probability values are obtained:

Controls/neurotics	$\chi^2 = 6.992$	$P = < 0.01$
Obsessionals/hysterics	$\chi^2 = 0.005$	$P = \text{N.S.}$
Anxiety states/hysterics	$\chi^2 = 2.298$	$P = > 0.10$
Anxiety states/obsessionals	$\chi^2 = 3.358$	$P = > 0.05$
Anxiety states/obsessionals + hysterics	$\chi^2 = 3.748$	$P = 0.05$

Mean movement intervals

Table 2 summarizes mean movement interval values and mean movement interval ranges for subjects who perceived ϕ . The total experimental group and the neurotic subgroups are seen to require a shorter interval between the stimuli in order to perceive movement, and they perceived movement between the samples was calculated control group. The significance of differences between the samples was calculated with a two-tail test of significance by means of the non-parametric Mann & Whitney U-test (Mann & Whitney, 1947). None of the differences in movement ranges was significant, but the following significant differences in mean movement intervals were found:

Controls/neurotics	$U = 322.5$	$P = 0.05$
Controls/anxiety states	$U = 144.0$	$P = 0.05$
Controls/obsessionals	$U = 101.0$	$P = 0.05$

with controls in each case having the larger mean movement intervals. Although the neurotic subgroups were not significantly different in terms of mean movement

intervals, those of the obsessional group tended more markedly towards objective simultaneity of the pair of stimuli than did the other neurotic groups.

Table 2. *Mean movement intervals and movement ranges in m/sec.*

	Controls	Neurotics	Anxiety states	Obsessionals	Conversion hysterics
Mean movement intervals					
$\bar{x}$	128.0	106.4	103.4	98.3	117.7
S.D.	63.1	51.8	52.1	55.4	46.7
Range	38-209	25-165	25-164	51-165	63-153
Mean movement ranges					
$\bar{x}$	49.1	44.3	47	38.8	47.1
S.D.	35.8	40.9	37.5	32.4	51.1
Range	0-124	0-128	0-128	0-82	0-112

Correlating mean movement interval with ambiguity avoidance scores on the previously described perceptual tests, only three significant correlations (expressed by Spearman's Rho corrected for ties by Kendall's (1948) formula) were found among the twelve that were calculated. These were with:

Weight discrimination	$\rho = 0.312$	$P = 0.05$
Length discrimination	$\rho = 0.284$	$P = 0.05$
Ambiguous pictures III (car-dog series)	$\rho = 0.294$	$P = 0.05$

Correlates of perception/imperception

The results so far have been concerned with the differences between control and experimental groups in the *perception* of apparent movement, differences in what roughly speaking are movement thresholds and perceptual correlates of these thresholds. To test Hypothesis II it was necessary to compare all those subjects perceiving phi with those who obtained no impression of movement on the perceptual tests enumerated above. Comparison of groups was again by non-parametric technique. The findings are summarized in Table 3.

The age parameter

Table 3 shows that subjects 'not seeing' phi were significantly older than those perceiving the phenomenon. While the former group consisted of 79 % neurotics and only 21 % normal controls, thus supporting Hypothesis II, it seemed necessary to partial out the effect of age. Since the total number of subjects was not big enough to yield a sizeable sample for a single-year or even a 5-year age range, a 10-year range had to be employed for purposes of comparison. Subjects in the 25-34 and 35-44 years ranges perceiving and not perceiving apparent movement in the experimental and control groups, were compared in their two groups separately. Table 4 shows that for both experimental and control groups perception of phi occurs significantly more frequently in the younger age groups. Comparing neurotics and controls for 'seeing' and 'not seeing' phi in the 35-44 years age range, no significant difference between the groups emerges.

Table 3. Comparing those 'seeing' phi and those 'not seeing' phi for differences on other tests

Samples: $m = 38, n = 64$.

For significance:

at 1 % level $U = 844$
at 5 % level $U = 933$ } on a two-tail test of significance

Other tests	U score	P
Weight discrimination	783	< 0.01
Length discrimination	831	< 0.01
Brightness discrimination	809	< 0.01
Block sorting test 'Can't decide'	986	N.S.
Block sorting test 'Others' score	1038	N.S.
Ambiguous pictures I	763	< 0.01
Ambiguous pictures II	827	< 0.01
Ambiguous pictures III	835	< 0.01
'Ambiguity' scale	841½	< 0.01
Reversible perspectives		
Autonomous reversals	897	< 0.05
Rubin figure	942½	> 0.05
Necker cube		
Controlled reversals	899	< 0.05
Rubin figure	806	< 0.01
Necker cube	1059	N.S.
Intelligence	719	< 0.01
Age		

Table 4. Perception of phi in two age groups

Age	Neurotics		Controls	
	'Seeing' (%)	'Not seeing' (%)	'Seeing' (%)	'Not seeing' (%)
25-34	69	31 $\chi^2 = 4.49$	88	12 $\chi^2 = 4.04$
35-44	37	63 $P = 0.03$	55	45 $P = < 0.05$

(χ^2 was calculated on numbers in each cell, not percentages.)

IV. DISCUSSION

The results may be taken to support both the hypotheses. The experimental group of neurotics obtained an impression of apparent movement significantly less frequently than did the control subjects, and required significantly shorter stimulus intervals for the perception of phi. Mainly anxious patients obtained an impression of movement significantly more frequently than did the combined groups of obsessionals and conversion hysterics. The combination of these latter two groups was considered justified in view of the very small difference between them, which was in any case anticipated on the basis of the earlier theoretical considerations. While the differences between the group of anxiety states on the one hand, and the separate groups of obsessionals and conversion hysterics on the other, were not significant, the data trends were at least in the predicted direction. The reason for this less satisfactory result may be sought in the degree of subjective uncertainty which was apparent in the patients with a primary symptomatology of anxiety: they were more obviously worried about giving a 'wrong' answer, and they may well have experienced the movement phenomenon but rejected it in case they would be thought silly. By

contrast, hysterics as well as obsessionals displayed a confident—if not over-confident—approach to the situation, and the question of doubt rarely arose, not even, surprisingly, among the obsessional sample. There was no evidence that they experienced and rejected the phenomenon.

While the number of significant correlations between mean movement intervals and the other measures of perceptual ambiguity and conflict exceeded chance, they are, of course, small. It may be considered surprising that the obtained relationships were found at all in view of the relatively primitive method employed to assess the 'best' movement interval. While the relationship obtained is somewhat tenuous, its presence was postulated and it ought to be considered in the general framework of findings reported here. Considerably greater importance may perhaps be attached to the results given in Table 3. The difference between subjects perceiving and not perceiving apparent movement on the other tests is marked and consistent. Subjects who failed to experience phi, defined in this context as an unstructured, indeterminate, perceptually unstable stimulus situation likely to involve responding in an ambiguous, perceptually conflictful setting, showed a significantly greater avoidance of the perceptual ambiguities of the test battery. Hypothesis II may therefore be considered supported, in that the findings are consistent with a theoretical framework in terms of generalized perceptual attitudes, functioning in this case as a defensive cognitive control against ambiguity and conflict aroused by ill-defined and unpredictable aspects of the environment. In view of the composition of the avoidance group, this control mechanism may be justifiably associated with neurotic defensive controls as has been previously postulated.

It has already been shown that the control and experimental groups were reasonably equated for sex and intelligence, and that performance on the tests of the battery, apart from the phi, was not materially related to age (Hamilton, 1956, 1957*b*). It is important to remember this when considering the found correlation of age and imperception of apparent movement. The data show that older subjects, whether normal controls or experimental neurotics, are less likely to perceive the phi phenomenon, but that even allowing for age neurotics found perception more difficult than the controls. When considering differences in mean movement interval for all subjects 'seeing' phi, no significant differences were found when comparing the two quartiles containing respectively the oldest and the youngest subjects. The nature of the relationship between imperception of phi presented by the single exposure method and age, whether in terms of total imperception or in terms of shorter stimulus intervals, is therefore not fully clarified by this experiment. It may be assumed, however, that the evidence for Hypothesis II is not controverted by the findings concerning the relationship with age.

While some support remains for Hypothesis I, despite the involvement of the age factor, a more serious qualification of the evidence for this hypothesis might be raised by suggesting that the results of the neurotic group may have been due to their being patients. It was not possible to test a control group of patients with physical disabilities, but this kind of alternative explanation would also have to account for the differences and similarities found between the *neurotic subsamples*.

A further shortcoming is the absence of direct evidence linking degree of rejection of the phi phenomenon with degree of neurosis. While such an investigation ought to be

done, it was found impracticable to test patients before treatment as well as after obtaining relief following it. Since valid and unidimensional quantitative measures of neurosis are not yet available, pre- and post-treatment ratings of degree of illness would have to be used in such an investigation.

If it were to be assumed that the most important factor in the present findings of imperception of *phi* was age, it would still not be possible to decide at present whether inability to perceive apparent movement in the older age groups was due to some lack of physiological facilitation, or to what is often and vaguely referred to as the greater 'rigidity' of older people. A physiological explanation at least for the present age groups seems contra-indicated by the absence of differences in mean movement intervals between the oldest and youngest subjects perceiving *phi*. While the possible relationship between any attitude specific to older age groups and avoidance of ambiguity was not directly investigated in this and the associated studies, age did not significantly correlate with the other tests. It is, therefore, cautiously assumed that total imperception of *phi* may involve a more fundamental rejection of perceptual instability and ambiguity than imperception of *phi* with longer stimulus intervals only. While the present study supports such an interpretation, a more analytic investigation in this area is clearly necessary for greater certainty. Not only would one have to look at the performance of very much older subjects, but investigation of subjects with arterio-sclerotic or other vascular disorders would be particularly useful for a theory of cortical facilitation in relation to the perception of *phi*. Within the present explanatory framework, however, total imperception is equated with a greater subjective need to avoid ambiguity and conflict.

I am indebted to Prof. C. A. Mace, Mr B. Foss and Dr J. Brown of Birkbeck College for methodological and statistical advice, to Mr E. Wasservogel for the construction of the apparatus, and to the Governors of Birkbeck College and the Central Research Fund of the University of London for financial support.

REFERENCES

- BENDER, M. & TEUBER, H. L. (1949). Disturbance in visual perception following cerebral lesions. *J. Psychol.* **28**, 223-33.
- BRENNER, M. W. (1956). The effects of brain damage on the perception of apparent movement. *J. Personality*, **25**, 202-11.
- FRENKEL-BRUNSWIK, ELSE (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Personality*, **18**, 108-43.
- GANTENBEIN, M. M. (1952). Recherche sur le développement de la perception du mouvement avec l'âge. *Arch. Psychol.* **33**, 198-294.
- HAMILTON, V. (1956). Reaction to uncertainty and ambiguity as a personality variable. Unpublished doctor's dissertation, University of London.
- HAMILTON, V. (1957a). Conflict avoidance in obsessionals and hysterics and the validity of the concept of dysthymia. *J. Ment. Sci.* **103**, 666-76.
- HAMILTON, V. (1957b). Perceptual and personality dynamics in reactions to ambiguity. *Brit. J. Psychol.* **48**, 200-15.
- JEEVES, M. A. & BRUNER, J. S. (1956). Directional information and apparent movement. *Brit. J. Psychol.* **8**, 107-13.
- JONES, E. E. & BRUNER, J. S. (1954). Expectancy in apparent visual movement. *Brit. J. Psychol.* **45**, 157-65.
- KENDALL, M. G. (1948). *Rank Correlation Method*. London: Griffin.

- KLEIN, G. S. (1951*a*). The personal world through perception. In *Perception—An Approach to personality*. Ed. Blake, R. R. and Ramsay, G. V. New York: Ronald.
- KLEIN, G. S. & SCHLESINGER, H. (1951*b*). Perceptual attitudes towards instability: 1. Prediction of apparent movement experiences from Rorschach responses. *J. Personality*, **19**, 289–302.
- KLEIN, G. S. & SMITH, G. J. W. (1953). Cognitive controls in serial behaviour patterns. *J. Personality*, **22**, 188–213.
- KLEIN, G. S. (1954). Need and regulation. In *Nebraska Symposium on Motivation*. Ed. Jones, M. R., Lincoln, University of Nebraska Press.
- MANN, H. B. & WHITNEY, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* **18**, 50–60.
- MEILI, R. & TOBLER, E. (1931). Les mouvements stroboscopiques chez les enfants. *Arch. Psychol.* **23**, 131–56.
- SHAPIRO, M. B. (1954). A preliminary investigation of the effects of continuous stimulation on the perception of apparent movement. *Brit. J. Psychol.* **45**, 58–67.
- WERNER, H. & THUMA, A. (1942). A deficiency in the perception of apparent motion in children with brain injury. *Amer. J. Psychol.* **55**, 58–67.

(Manuscript received 6 June 1959)

A STUDY OF PREFERENCES IN FEMININE BEAUTY

By A. H. ILIFFE

University College of North Staffordshire

A survey of previous studies of aesthetic preferences shows little attention to examples of natural beauty and none to preferences in the field of human beauty. In inquiries using factor analysis it is necessary to distinguish different uses of the term 'the general factor in aesthetic judgements'. In the present inquiry twelve photographs of women's faces taken in uniform conditions were inter-ranked for 'prettiness' by 4355 readers of a national daily newspaper. Twenty-two inter-correlations between mean rankings of men and women, of four age groups, and of six regional groups were each of the order of 0.9. Conformity of opinion within groups was significantly different only for the 55+ age group who showed less tendency to agree than other age groups. Factor analysis of a small random group of individual rankings resulted in a general factor accounting for 43 % of the variance, and a small bipolar factor. Comparison of mean rankings of representative samples of five occupational groups showed that social status has little influence on preferences. Limitations of the technique of inquiry are discussed and the results are considered in the light of the hypothesis that norms of human beauty are culturally determined.

I. INTRODUCTION

Previous studies of aesthetic preferences have tended to follow one of two methods of inquiry, and have been concerned with a fairly limited range of material. The two methods consist equally in presenting to a group of subjects a set of objects of some aesthetic interest; in the first method subjects are asked to arrange these in order of preference; in the second they are required to rate each of the objects in terms of a number of given characteristics (e.g. emotional appeal, sentimentality). Most investigators have used as material what may broadly be called 'works of art'. Reproductions of paintings and/or sculptures were used by Burt (1924), Dewar (1938), Eysenck (1940*a*, 1941*a*), Peel (1945) and Pickford (1948). Pickford also took his subjects to see actual paintings in a gallery, and McElroy (1950) recorded the preferences of schoolchildren to sculptures in an open-air exhibition. Poems and passages of prose were used by Williams, Winter & Wood (1938), and poems by Eysenck (1940*b*), Gunn (1951) and Britton (1954); recorded musical works by Pickford (1948). Burt & Bulley (1933) used photographs of furniture, glass- and silver-ware, printed fabrics, embroidery and jewellery; Eysenck (1940*a*) included photographs of bookbindings, vases and embroidery in his material.

Outside the field of works of art (broadly interpreted) most attention has been paid to preferences for geometric shapes or line-drawings (Fechner, 1876; Birkhoff, 1933; Nienstedt, 1951; McElroy, 1954) and for single colours (see Eysenck (1941*b*) for a summary). It appears that almost no investigation has been made of preferences for examples of natural beauty; Eysenck included one set of landscape photographs in his demonstration of the 'K'-factor (Eysenck, 1941*a*), and in his earlier paper (Eysenck, 1940*a*) obtained rankings of a set of 'Blotters' which consisted of 'coloured photographs of "Bathing Girls"'. Though no details of instructions to subjects, nor of subjects' comments, are given, it is likely that these pictures, in common with other sets used in the same inquiry, were ranked on the general appearance of each picture, including colouring and composition, and that the relative attractiveness of each girl portrayed played a comparatively minor part.

II. THE 'GENERAL FACTOR' IN AESTHETIC PREFERENCES

A familiar feature of several of these investigations is that the orders of preference of individual subjects tend to correspond with each other to an extent which is highly unlikely to be due to chance. These findings are sometimes loosely referred to as indicating the existence of 'a general factor of aesthetic judgement' or 'capacity' (e.g. Pickford, 1956). It is important to distinguish between (a) general factors which result from the analysis of the inter-correlations of orders of preference (as in Dewar, 1938); (b) those which derive from the analysis of inter-correlations of totals of ratings for various characteristics (as in Pickford, 1948, and Gunn, 1951); and (c) Eysenck's general factor which is derived from inter-correlations of coefficients of conformity. General factors of type (a) state the extent to which individual orders of preference tend to correspond. General factors of type (b) state the extent to which pictures (for example) rated high on each characteristic (e.g. 'aesthetic design', 'emotional expression') tend to be rated high on all characteristics. Such general factors could be the result of the all-or-none response of indiscriminating subjects, rating pictures they liked high in all characteristics. The possibility that this 'halo effect' has operated would seem to justify caution in awarding this general factor the title of the 'Aesthetic Factor' (Pickford, 1948). Eysenck's statistical treatment begins with the calculation for each subject of a coefficient which indicates his tendency to conform to the average ranking for each set of pictures. These coefficients of conformity are treated in exactly the same way as scores on tests. When the lists of coefficients (subjects' 'scores') are correlated between one set of pictures and another (tests) a matrix is produced which yields a general factor. This states the extent to which subjects' tendency to conform in any one set of pictures approximates to his over-all tendency to conform. Eysenck's bipolar factor was derived from a matrix of inter-correlations of orders of preference.

The 'generality' of any factor is limited to the actual subjects tested and the material used; the existence of corresponding factors in larger populations can only be inferred if the samples, of persons or of material, are representative.

III. DESIGN OF THE INVESTIGATION

The present inquiry was planned to extend the investigation of aesthetic preferences to a field which had not previously been the object of this type of inquiry, namely feminine beauty. The original preparations, which were on an extremely modest scale, came somewhat fortuitously to the attention of one of the national daily papers, who suggested that their readers might be interested in taking part, and offered assistance in producing the experimental material and in tabulating the eventual results.

The material consisted of twelve black-and-white photographs of the faces of young women selected from a very large number which were taken for this purpose by a staff photographer of the newspaper. The twelve photographs were chosen to provide as great a range as possible of types (i.e. in shape of face, colouring, shape of eyes, etc.). The following steps were taken to ensure the maximum possible uniformity of conditions which might affect the responses of subjects: (i) all the young women

portrayed were in the age range 20 to 25; none was a professional 'beauty' (i.e. film actress or model); (ii) all photographs were three-quarter profile showing head and shoulders; (iii) no clothing, jewellery or any kind of hair-ornament was visible; (iv) make-up was conventional; (v) all photographs were taken with identical lighting and by the same photographer. The possibility that some of the girls portrayed would be familiar to some readers who sent in their orders of preference was met by including on the published form a request 'if you think you recognise any of these girls please write her name here'; any form in which the space was filled in, whether correctly or otherwise, was excluded from the results.

The twelve photographs appeared in two sets of six in successive issues of the paper. Readers were asked to record on a published form their orders of preference for all twelve, and to indicate their age group, occupation and address. The 4355 valid returns were transferred to punched cards and sorted in forty-eight groups provided by the three classifications: men/women, four age groups, and six regions (see Appendix 1). The rank orders of preference for each of the twelve photographs were totalled for each of the forty-eight groups and the grand totals were computed for men and women separately, for each age group and for each region.

IV. RESULTS

(a) Orders of preference according to sex, age and region

When the proportions of men and women in the various age groups who submitted returns were compared with corresponding proportions in the figures of the 1956 census of the United Kingdom there was found to be a noticeable bias ($\chi^2 = 351.1$, $P < 0.01$) particularly in excess of women in the age group 35-54 (Table 1).

Table 1. *Comparison of returns with a proportionate sample of the population of the United Kingdom*

Age group	Returns		Sample	
	Men	Women	Men	Women
15-24	220	197	351	359
25-34	457	401	391	394
35-54	740	1136	783	817
55+	342	862	532	728
Sub totals	1759	2596	2057	2298
Total		4355		4355

In calculating the over-all orders of preference the totals for men and women in four age groups were weighted to correct this bias; similarly, in calculating the grand totals for men and women separately totals for separate age groups were weighted. The average orders of preference which resulted showed very close correspondence between men and women, between different age groups and between people from different regions. Table 2 shows the rank correlation coefficients between the average orders of preference for different groups, and between each group and the general order of preference.

Table 2. *Correlations between orders of preference*

	Men		Women	General
	Men	Women	0.88	0.93
			—	0.99
Age groups				
Age group	25-34	35-54	55+	General
15-24	0.98	0.94	0.80	0.96
25-34	—	0.98	0.85	0.99
35-54	—	—	0.90	0.99
55+	—	—	—	0.91

(b) *Conformity of opinion within groups*

Calculation of the variance of the mean rank orders of preference in each group (σ_{MR}^2) gives an indication of the degree to which people within that group agree in their judgements; if the rankings of all subjects in a group are identical $\sigma_{MR}^2 = 11.91$; if rankings are random $\sigma_{MR}^2 = 0$. Table 3 shows the variances for men and women for each age group and for each region; the over 55 age groups show a significantly smaller variance (i.e. less correspondence of opinion) than other age groups.

Table 3. *Comparison of extent of agreement within groups.*
Variance of mean rank orders (σ_{MR}^2)

Men	3.6	Regions	
Women	3.5	Wales	2.7
Age groups		North	3.0
		West	2.8
	15-24	Midlands	3.4
	25-34	East	3.3
	35-54	South	3.0
55+	2.5		

With this number of subjects there is clearly no possibility of calculating correlations between persons and using factorial analysis. A sample of twelve was selected on a random basis, and rank correlations between individual orders of preference were calculated. Of the sixty-six coefficients only five were negative, and none of these was significant. Thirteen of the positive coefficients were greater than 0.6; the mean of all sixty-six coefficients was 0.35. Analysis by Thurstone's centroid method produced a general factor with saturations ranging from 0.36 to 0.88 accounting for 43% of the variance and a bipolar factor accounting for a further 16%. Saturations on neither factor showed any significant correlation with sex or age of the subject.

(c) *The influence of occupation*

Although subjects had been asked to record their occupation this information was not coded for the mechanical tabulation. In order to gain some evidence of the effect of occupational status on order of preference a representative sample of fifty

was selected for each of the five classifications of occupation in the Registrar General's scheme, i.e.:

Class definition	Examples
I. Professional, etc. occupations	Architect, surgeon
II. Intermediate occupations	Teacher, sales manager
III. Skilled occupations	Insurance clerk, engine driver
IV. Partly skilled occupations	Bus conductor, fitter's mate
V. Unskilled occupations	Messenger, kitchen hand

The occupations recorded by subjects were classified by reference to the classification of occupations compiled for the 1951 census; returns on which vague terms were used or insufficient detail was given were excluded from the samples. Each sample of fifty contained numbers of men and women in the four age groups which were approximately proportional to the numbers recorded in the 1956 census of the United Kingdom with the result that a comparison of the average rank orders for each sample was freed from any possible bias due to age or sex of subjects in the sample. Inter-correlations between the average orders of preference for the five occupational classes are shown in Table 4. As will be seen, there is no marked tendency for the opinions of people in any one occupational class to resemble more closely those of people in adjacent occupational classes.

Table 4. *Correlations between orders of preference: occupational groups*

Occupational class	II	III	IV	V	σ^2_{MR}
I	0.96	0.95	0.91	0.80	3.69
II	—	0.94	0.93	0.78	4.39
III	—	—	0.94	0.88	2.53
IV	—	—	—	0.76	2.53
V	—	—	—	—	1.82

An over-all order of preference based on these five samples differs very little from the general order obtained from all returns ($\rho = 0.99$).

From these results it appears that any tendency for the 4355 subjects to be disproportionately distributed among occupational classes, which might arise from the complexion of the newspaper of which they are all readers, or from differences in interest and readiness to take part in an investigation of this kind, is likely to have very little bearing on the general order of preference.

V. DISCUSSION

Writers on aesthetics have made little reference to the beauty of the human face. A popular view is that ideas of facial beauty vary widely from one person to another. Before discussing the relevance of the present findings, however, two reservations must be made.

We are dealing here with statements of preference for photographs, which lack the colour, the mobility, and to an important extent the expressiveness of living faces. Had it been possible to include these aspects, by the use of colour films, for example, the number of criteria on which subjects could have based their judgements

would have been increased, and a greater disparity of opinion might have resulted. A further limitation lies in the use of terms: the appraisal of human beauty has an extensive vocabulary, and as well as 'beautiful' the words 'pretty', 'handsome', 'good-looking', 'attractive', 'charming' (and many others) are widely used. It is frequently claimed that each of these terms can be differentiated, by denotation if not by connotation. The form on which subjects in this inquiry recorded their preferences made use of the term 'pretty', and it is possible that the substitution of another word might have produced different orders of preference.

When these reservations are given the fullest possible weight, however, the results of the inquiry remain striking, and suggest that a common basis of judgement is shared by men and women of all ages and in all parts of the country. The effect of differences in technical excellence or in familiarity, which may tend to produce a consensus of opinion when preferences for works of art are explored, are ruled out here by the nature of the material and the conditions under which the photographs were taken. Again, it would undoubtedly be possible to demonstrate wide agreement in orders of preference if the set of faces were to include both examples of highly conventional beauty and faces with obvious blemishes, but in fact the extent to which the faces used varied in good looks was deliberately kept very narrow, and included none which could be called (at least as far as the experimenter could surmise) either 'very beautiful' or 'plain'.

The question then arises: what is the origin of this common basis of judgement, for although the correlations which have been demonstrated were between the summated orders of judgement for large groups of subjects it seems justifiable to infer, with the support of the results of the factor analysis of the small random sample, that a general factor would emerge from the analysis of the whole of the data. One possible answer to this question is that human faces possess in varying degrees some intrinsic characteristic (e.g. of 'harmony' or 'balance') which is shared by all beautiful things. A more likely explanation is that culturally determined norms are transmitted through education (and perhaps through familiarity with the visual and dramatic arts); some slight corroboration for this hypothesis may be derived from the decreasing extent of agreement among subjects in lower occupational classes (see Table 4). Such norms would be susceptible to slow processes of change, and in this connexion it may be relevant to refer to the lower correlations between the order of preference of subjects in the 55+ age group and those of younger subjects (see Table 3).

These surmises must of course be extremely tentative. More intensive research is needed to produce more clear-cut evidence, and to explore the extent to which the agreement demonstrated among subjects in the United Kingdom would extend to people of other cultures.

APPENDIX 1. DEFINITION OF REGIONS

Region 1	Wales	
Region 2	North	Scotland; and Cumberland, Durham, Lancashire, Yorkshire, Northumberland, Westmorland
Region 3	West	Cornwall, Devon, Dorset, Gloucestershire, Herefordshire, Monmouthshire, Shropshire, Somerset, Worcestershire
Region 4	Midlands	Cheshire, Derbyshire, Leicestershire, Lincolnshire, Northamptonshire, Rutlandshire, Staffordshire, Warwickshire

Region 5	East	Bedfordshire, Cambridgeshire, Essex Hertfordshire, Huntingdonshire, Norfolk, Suffolk
Region 6	South	London, Berkshire, Buckinghamshire, Oxfordshire, Surrey, Hampshire, Sussex, Kent, Middlesex, Wiltshire

The invaluable assistance of the staff of the *News Chronicle-Dispatch* is gratefully acknowledged, and particularly that of Mr Charles Widdup and Mr Barnet Saidman.

REFERENCES

- BIRKHOFF, G. D. (1933). *Aesthetic Measure*. Cambridge, Mass.: Harvard University Press.
- BRITTON, J. N. (1954). Evidence of improvement in poetic judgement. *Brit. J. Psychol.* **45**, 196-208.
- BURT, C. (1924). Psychological tests of educable capacity. In *Report of the Consultative Committee of the Board of Education*. London.
- BURT, C. & BULLEY, M. (1933). Articles in *The Listener*: 18. i. 33, 8. ii. 33, 27. xii. 33. London: B.B.C.
- DEWAR, H. (1938). A comparison of tests of artistic appreciation. *Brit. J. Educ. Psychol.* **8**, 29-49.
- EYSENCK, H. J. (1940a). The general factor in aesthetic judgements. *Brit. J. Psychol.* **31**, 94-102.
- EYSENCK, H. J. (1940b). Some factors in the appreciation of poetry and their relation to temperamental qualities. *Character & Pers.* **9**, 160-7.
- EYSENCK, H. J. (1941a). 'Type'-factors in aesthetic judgements. *Brit. J. Psychol.* **31**, 262-70.
- EYSENCK, H. J. (1941b). A critical and experimental study of colour preferences. *Amer. J. Psychol.* **54**, 385-94.
- FECHNER, G. T. (1876). *Vorschule der Aesthetik*. Leipzig: Breitkopf and Härtel.
- GORDON, D. A. (1955). The artistic excellence of oil paintings as judged by experts and laymen. *J. Educ. Res.* **48**, 579-88.
- GUNN, D. G. (1951). Factors in the appreciation of poetry. *Brit. J. Educ. Psychol.* **21**, 96-104.
- McELROY, W. A. (1950). Responses to traditional and modern sculpture and factors influencing its recall. *Quart. Bull. Brit. Psychol. Soc.* **1**, 310-16.
- McELROY, W. A. (1954). A sex difference in preferences for shape. *Brit. J. Psychol.* **45**, 209-16.
- NIENSTEDT, C. W. (1951). Preferences for rectangular proportions in college students and the aged. *J. genet. Psychol.* **78**, 153-8.
- PEEL, E. A. (1945). On identifying aesthetic types. *Brit. J. Psychol.* **35**, 61-9.
- PICKFORD, R. W. (1948). 'Aesthetic' and 'technical' factors in artistic appreciation. *Brit. J. Psychol.* **38**, 135-41.
- PICKFORD, R. W. (1956). Factorial studies of aesthetic judgements. In ROBACK, A. A., *Present-Day Psychology*. London: Peter Owen.
- WILLIAMS, E. D., WINTER, L. & WOOD, J. M. (1938). Tests of Literary Appreciation. *Brit. J. Educ. Psychol.* **8**, 265-84.

(Manuscript received 31 August 1959)

PUBLICATIONS RECENTLY RECEIVED

Plans and the Structure of Behavior. By GEORGE A. MILLER, EUGENE GALANTER and KARL H. PRIBRAM. New York: Henry Holt. 1960. Pp. x + 226. \$5.00.

'If a man is like a computer, then the man must have somewhere an organized set of instructions that he attempts to execute. That is to say, the man must have a Plan' (p. 197). By taking the analogy between man and computer seriously, we are driven to search for the plans people have and to tease out their characteristics. This book describes that search as conducted by three already reputable men during a year spent together in California at the Centre for Advanced Study in the Behavioral Sciences. The trio would be the last to claim that their search is complete; they modestly present their book as their 'intellectual diary for the year' and hope that its 'spontaneous, argumentative, personal style' will not prevent our recognizing that some of their ideas are too good to lose. For the authors, the recent efforts to construct computers which simulate psychological processes mean that 'perhaps some of the more fanciful conjectures of the "mentalists" should now be seriously reconsidered' and that 'psychologists have been issued with a new license to conjecture' (p. 56). And conjecture is what the authors do. But if unconventional in places the conjectures are never wild, nor do they ever become so elaborate as to mask the psychological phenomena they were produced to illumine. There is nothing blinkered, unduly technical nor, in a narrow sense, materialistic about this book—the authors label themselves 'subjective behaviourists'—and it will certainly be cited and quoted for some time to come.

The authors devote the first three chapters to elaborating their theoretical apparatus, their spectacles through which, in the remaining chapters, they propose to look anew at a variety of psychological events. In chapter 1, they announce that their preferences are all on the side of cognitive theorists (e.g. Bartlett, Kohler, Tolman), express their concern about filling the theoretical gap between cognition (or, as they term it, Image: 'all the accumulated, organized knowledge that the organism has about itself and its world') and action, and then they suggest that what is needed to fill this gap is the notion of Plan. 'A Plan is any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed' (p. 16). Thus Plans refer to the complete specifications of, or programmes for, every detailed operation in courses of action; they also refer to more or less sketchy outlines of courses of action yet to be performed, to what we colloquially call our plans, to what William James described as 'these rapid premonitory perspective views of schemes of thought not yet articulate'. At a relatively molar level Plans may be called strategies; at a relatively molecular level, tactics. The generation and execution of Plans are such commonplace aspects of behaviour that it is startling to realize how neglected they have been by psychologists. However, the authors do more than merely draw attention to this neglect; they propose, through their notion of 'hierarchy', the general nature of the organization that is a Plan. In chapter 2, a damaging blow is delivered at the reflex as the unit of behaviour and the TOTE is presented as the elementary unit of psychological analysis. TOTE stands for Test-Operate-Test-Exit. There is representation of an anticipated state and a present state; the incongruity between these two states is the Test which leads to some activity, the Operation, and to a second Test of congruity between the anticipated and (now altered) present state; if there is congruity, then Exit takes place and behaviour control shifts elsewhere; if there is still incongruity, a second Operation is carried out. It could fairly be said that TOTE is a rephrasing of the trial-and-check unit so familiar in the literature of problem solving. But TOTE is conceived to be more general than this. 'The present authors feel that the TOTE unit, which incorporates the important notion of feedback, is an explanation of behaviour in general, and of reflex action in particular' (p. 29). The TOTE unit operates at all levels on the molar-molecular continuum and, through a discussion of nail-hammering activity, the authors illustrate how several TOTE units at different levels may form a compound hierarchical TOTE unit or Plan which issues in a sequence of activities. Chapter 3 rounds off the authors' construction of their theoretical apparatus and compares it with the construction of

other theorists. The spectacles are now ready for wear and, through them, the authors proceed to look at the confusing world of behaviour and experience. In looking with them we find that their spectacles throw many matters into clearer focus without distortion. Some promising avenues for research are opened up. The TOTE unit proves to be a powerful conceptual device, simple yet complex; so does its compound, the Plan, with its hierarchical-cum-sequential structure: it is, in essence, the able handling of hierarchy and of sequence that is the book's great merit. And, throughout, there is freshness and excitement. It cannot be said that the basic conceptual apparatus is entirely novel nor that it does not reflect the theoretical climate of the times. Something like it appears in the writings of several psychologists who are influenced by communication and computer theory, also in those of some psychologists who either wrote before the behaviourist revolution or were unaffected by the excesses of this historical movement, e.g. Woodworth with his conceptualization of learning in terms of a question-and-answer process, or Harlow with his conception of 'organised response pattern'. However, it can be said that the notions basic to this book (and the reviewer regards them as notions of immense potentiality) have never before been presented so explicitly nor so capably.

Once equipped with their theoretical spectacles, the shifting direction of the authors' regard may be summarized by listing the titles of the book's successive chapters. In *Values, Intentions, and the Execution of Plans* some notions of Lewin are dusted off and clarified. *Instincts* are briefly viewed as inherited Plans. *Motor Skills and Habits* are looked at as hierarchical organizations and there is a valuable commentary on the relations between mastering a skill (landing an aircraft is the main example used) and having a verbal description of the skill. *The Integration of Plans* within both the individual and the social group is followed by a discussion of *Relinquishing the Plan* in which we are given a novel and suggestive account of hypnosis. Individual and cultural differences in Plans are indicated in *Nondynamic Aspects of Personality*, e.g. their source, span, detail, flexibility, coordination, and location and kind of stop-order. In *Plans of Remembering* the often overlooked point is made that the memorizer's activity is largely devoted to constructing a Plan for recall. And *Plans for Speaking* brings the authors' viewing to a climax with an impressive analysis of the 'grammar Plan' guiding the generation of a spoken sentence. It is convincingly argued that this plan cannot be of a 'Markovian', left-to-right sort and that it must be of a hierarchical kind which 'can settle the future and past simultaneously and leave the growing tip of the sentence in the middle, rather than at the extreme right' (p. 151). This discussion illumines not only the serial ordering of language but the ordering of much activity besides and leads the authors to comment that 'the capacity to use Plans to construct Plans to guide behavior' (p. 157) may be 'the key to an evolutionary breakthrough as important as the development of lungs and legs' (p. 158). The next two chapters are concerned with problem solving and the merit of the view that this may be regarded as a search activity. *Plans for Searching and Solving* makes good use of Bruner's experimental work on strategies while *The Formation of Plans* takes its cue from Polya's strangely neglected writings on heuristic devices in problem solving and points out some of the ways in which computers have been designed to perform complicated information-processing according to heuristic principles. The last chapter provides *Some Neuropsychological Speculations* and its closing words convey something of the flavour of the entire book. '... The notions of a reflex telephone system with an enigmatic switch-board, or inhibitions and excitations rippling majestically over the surface of the brain, or little homunculi inside the pineal glands of little homunculi inside the pineal glands of little homunculi *ad infinitum*, or empty black boxes that absorb S's and emit R's, are so thoroughly unsatisfactory that, although the present ideas may be wrong, they are likely to be a great deal less wrong than the metaphors many psychologists have used heretofore. Anybody who tries to do the research needed to put this approach to test will discover things that he would not otherwise have thought to look for' (p. 209). The reviewer need say no more than to express his complete agreement with the last quoted sentence and his gratitude for having been allowed to wear theoretical spectacles with such clarifying properties: it matters little if these particular spectacles are ultimately discarded for, even so, they could hardly have failed to enrich our psychological vision.

I. M. L. HUNTER

Explanation of Human Behaviour. By F. V. SMITH. London: Constable. 1960 (2nd edition). Pp. xii + 460. 30s.

The first edition of Professor Smith's book appeared in 1951. Nine years later he has produced a second edition dropping 'The' from the title (formerly '*The Explanation of Human Behaviour*') and adding 184 pages to the text. He has given us a greatly improved book, not only larger, but a more mature and 'softer' Smith.

The old Smith was markedly Hullian in sympathy. Hull's system was singled out for special commendation, and regarded as a model, apart from minor blemishes, of what a psychological system should be. The new Smith is much more guarded. He does not consider that the whole of Psychology can be fitted into a system of the sort Hull constructed. 'Hull had actually attempted more than the resources of the subject permitted; but it was a courageous attempt which had many beneficial results.'

The new Smith is more sympathetic to what Hebb has called the 'left wing' movements in psychology, Gestalt Psychology and Psychoanalysis. Indeed the major change in the new edition, accounting for 106 of the 184 new pages, is the addition of two new chapters on 'The Gestalt Approach', and 'The Freudian Approach' respectively. The chapter of Gestalt is a very useful exposition and criticism of the Gestalt position, which could hardly be bettered. The chapter on Freud is sympathetic and balanced. These two chapters undoubtedly serve to correct a certain one-sidedness in the first edition, though we must remember that the author never set out to provide a complete picture of recent psychology.

Smith has not been afraid to change his views or to correct his mistakes. The old chapter on McDougall was marred by serious limitations of outlook. The new chapter has been extensively re-written and brought into line with our present understanding of instinctive mechanisms. The early chapters of the book on the nature of explanation have been considerably revised, enriched with examples of recent research and expressed less dogmatically. Instead of laying down 'the requirements for adequate explanation' and determining 'how certain systems in the literature of Psychology meet these requirements' (1st ed.) Smith discusses 'what can be expected of a system', and 'in appraising any system some important considerations... would appear to be...' (2nd ed.). This is certainly a 'softer' Smith, who looking back at the recent history of Psychology concedes 'that an era of system building would appear to have passed'.

No doubt this revision will give Smith's book a new lease of life. Combining as it does a critical appreciation of theory, historical perspective, and a rich diet of empirical findings, it is a valuable book for the honours student in psychology.

L. S. HEARNshaw

Personality and Persuasibility. By IRVING L. JANIS, CARL I. HOVLAND, PETER B. FIELD, HARRIET LINTON, ELAINE GRAHAM, ARTHUR R. COHEN, DONALD RIFE, ROBERT P. ABELSON, GERALD S. LESSER and BERT T. KING. New Haven, Conn.: Yale University Press. 1959. Pp. xiv + 333. \$10.00

This book is the second of a series of Monographs published by the Yale Communication and Attitude Change Program. The first of these, 'The Order of Presentation in Persuasion' appeared in 1957 and this second Monograph is similar to its predecessor in type. It consists of a number of separate chapters, written by different authors (there are ten of these listed on the title page), and each chapter reports an independent investigation or a separate theoretical analysis. Some of this work had been published before in journal articles or theses, and is here reprinted with minor changes. This form of presentation has obvious drawbacks; there is a good deal of duplication, and what might have been the most logical order of presentation cannot usually be observed because it would cut across the necessity of presenting each paper as an independent publication. Another disadvantage is the failure of the contributions to present a uniform level of excellence; instead the quality clearly follows the normal curve of distribution, with a rather high variance. To put it slightly differently, some of the contributions are outstanding, while some are so poor as to raise doubts about the appropriateness of including them in the volume. Essentially, the book deals with the concept of 'general persuasibility', i.e. 'a person's readiness to accept social influence from others irrespective of what he knows about the communicator or

what it is that the communicator is advocating'. Research on this problem had been begun by Ferguson and others as long ago as 1944, with positive results, but there is no doubt that the studies reported here put the concept on a much more secure foundation, establishing it as being of much wider significance than had previously been thought. The problem obviously is a factor of a fairly marked factor of 'persuasibility'. Their method is that of presenting a variety of attitude statements to their subjects for endorsement; these are then followed by a series of 'communications' regarding the topics dealt with by the attitude questionnaire. A second administration of the attitude questionnaire then ascertains the degree of change produced by the first, is then administered, after a suitable interval of time, and the attitude questionnaire is filled in a third time. Under these conditions it is found that subjects show a general tendency either to change their attitudes in line with the communication, or not to change them, regardless of the topic in question. A few subjects change their attitudes in the direction opposite to that advocated by the communication source. Slightly different methods from this were used for young children, but the essential paradigm remains the same.

From this beginning the work of the Hovland group extended along two main lines. In the first place, personality correlates of persuasibility were investigated. These included objective tests such as responses to social influence in an autokinetic situation, the tilting room/tilting chair test, and the Gottschaldt embedded figures test; they also included questionnaires and projective tests. Many of the correlations reported are significant and suggest that persuasibility is indeed related to certain aspects of personality. The second type of work reported dealt with developmental aspects of persuasibility in children and relates to such factors as child rearing practices and so forth.

It would be impossible to summarize the findings in detail, except to say that they are frequently significant and equally frequently contradictory. Thus, throughout the book, results for male subjects differ from those for female subjects; the authors make some attempts to explain this disagreement, but finally leave it unexplained. Other disagreements relate to age; correlations found at one age, may disappear or be inverted, at another age. This is not, of course, a criticism of the work; there is no reason why the same factors should work in the same direction for the two sexes, or at different ages. Hovland rightly draws attention to the complexity of the field and does not pretend to have achieved a premature unification.

Nevertheless, there are one or two criticisms that must be made. In the first place, results are sometimes taken at face value without regard to possible criticisms. Thus, the *F* scale is used, but no mention is made in the interpretation of the data of the recent work on *response set* which has completely altered the meaning of scores derived from this scale. Again, the statistical treatment is not entirely satisfactory. Large numbers of correlations or *t* tests may be presented in one table, and the significance value of each given as if it had been the only test actually made. This tends to give rise to spurious significance, and some correction should be made for the fact that each comparison is not the only one undertaken for that particular experimental group. More important perhaps than these criticisms is a rather more fundamental one. The authors constantly interpret the correlation between personality traits in parents and personality traits in children as being the causal product of the former. This, of course, is possible, but by no means certain. It is equally plausible to argue that genetic factors inherited by the children from the parents produced both parental and filial behaviour patterns, and thus gave rise to the observed correlations. It is, of course, customary in American writings in this field to disregard constitutional and hereditary factors, but from experimentalists of such ability one might have expected at least a slight *caveat* rather than complete disregard of alternative hypotheses to the environmentalist one.

H. J. EYSENCK

Humanism and Moral Theory. By R. OSBORN. London: Allen and Unwin. 1959.
Pp. 115. 18s.

Mr Osborn's aim, in his own words, is to 'look for an interpretation of ethics in humanist terms, an interpretation which does not go beyond the bounds of human life for its criteria of good and bad, right and wrong, yet avoids the difficulties of relativism that the rejection of absolute, extra-human standards seems to involve... to see, in other words, whether we can get something

of the meaning of such terms as "good" and "right" from a consideration of the psychological and social conditions in which these terms come to be used'.

His book is stronger on the psychological than on the philosophical side. He does not, except by implication, answer the arguments of theologians like Hastings Rashdall, who claim to demonstrate that there can be no valid morality without supernatural sanctions. But he gives a valuable account, from the psychological angle, of the development of moral concepts and moral attitudes: and, drawing heavily on Piaget, he suggests an illuminating parallel between the development from religious to humanist ethics, and the development from the authoritarian, superego morality of early childhood, when 'being good' means obeying the arbitrary injunctions and prohibitions of adults, to the socialized ego morality of later years, when to act right means to act for the common good.

The later, more philosophical chapters are concerned mainly with ethical theory, including such time-honoured topics as the relation between 'good', 'right' and 'ought'. These chapters lack authority and distinction: they give the impression that they have not been fully thought through, nor clearly related to the main theme. But none of these strictures applies to the first part of the book, which is stimulating and original. No proselytizing humanist can afford to miss chapter 1.

M. KNIGHT

The Social Desirability Variable in Personality Assessment and Research. By ALLEN L. EDWARDS. New York: Dryden Press. 1957. Pp. viii + 108. \$3.25.

In this slender, elegantly produced monograph, Prof. Edwards expounds, with possibly excessive lucidity, a series of studies tracking a variable likely to prejudice some personality research. Every statement on a self-assessment inventory refers to behaviour or feelings which can be rated on a dimension of social desirability (SD). More important, from the ratings so obtained can be predicted the proportion of endorsements the statement will obtain when used in a check-list purporting to measure anxiety, dominance, dependence or what have you: the higher the SD value, the more frequently will it be chosen. The SD assessments remain essentially the same even when the judges are psychotic patients or skid row alcoholics. Scoring on an SD scale correlates substantially with several well-known personality inventories and Q-sorts.

Of course, both the high agreement among the judges and the frequency of endorsing socially desirable items can be regarded as a tribute to the socialization agencies of society. It is to be expected, in any case, that behaviour selected for assessment by inventory compilers would have relevance to what common opinion finds troublesome. Clearly, Edwards's opinion is that the tendency to select the high SD items is a contaminating one. He suggests, for example, that the association between low Anxiety scoring and efficiency of learning can be reinterpreted by taking account of this tendency. It is the low scorers who are truly anxious—about public opinion—and it is the desire to please that makes them do well in the experimental task. The people who get high scores are those who are unanxious, in the sense that they do not trouble to conceal their weaknesses. Are the neurotic people, as defined by inventories, the only mentally healthy ones? Is truthfulness—not self-disparagement—the mark of the depressive? And how were all these inventories validated in the first place? Although Edwards claims that the tendency to select the desirable items is a stable personality trait, he has not chosen to pursue it, and thus answer some of these questions, by studying individual behaviour, its consistency over inventories measuring different traits, and the characteristics of the non-conforming judges, if any. Instead, he has devised an instrument to by-pass the problem, by offering choices between items paired for SD rating. This test is not included in the present book.

A systematic error occurs in the index. To the pagination of all references to the bibliography, 2 must be added. It is disconcerting to look up Eysenck, for example, and be directed to a blank page.

THELMA VENESS

Art and Illusion: a study in the Psychology of Pictorial Representation. By E. H. GOMBRICH. London: Phaidon. 1960. Pp. 446. 70s.

The widespread belief that all art is a reflexion of the optical world has not been greatly disturbed by the impact of Modern Art; nor has the study of the art and artifacts of primitive civilizations had much influence on the basic aesthetic assumptions of a substantial section of the educated

public; still less have the findings of modern psychology (notably) the work of Thouless, despite the critical approval they have achieved among the apologists of modern painting and sculpture. The Renaissance ideal of optical truth is still considered a valid criterion of aesthetic excellence. Yet a respectable body of research has shown beyond doubt that any attempt to copy exactly the appearance of the external world is not only not an artistic activity but not even possible. The resistance to the evidence of scientific research in the field of art is an interesting psychological phenomenon in itself which has not escaped notice. But are we right in assuming as many critics do that representational accuracy has nothing to do with the processes of art? In view of the present 'abstractionist dilemma' Prof. Gombrich's book is of the utmost interest and importance. His view is that jettisoning the discoveries and effects of representation in favour of 'non-figurative' art we may be in imminent danger of losing contact with the great masters.

Art and Illusion (the A. W. Mellon lectures in the Fine Arts 1956, National Gallery of Art, Washington) sets out to examine the psychological problem of pictorial representation; but the psychologist who expects a rigorous examination of the subject, together with a set of clear-cut conclusions, will be disappointed; not through any fault of the author (who is unnecessarily modest about his psychological status) but because no final conclusions are as yet possible in this field of psychology. Indeed, not the least merit of the book, apart from its thorough-going historical analysis of the problem and the massive evidence adduced, lies less in its stimulating views than in the range and type of question it asks and the numerous incidental problems it pinpoints. Timeously published, the book will remain a valuable work of reference for a long time to come.

A study so vast in its range and so packed with out-of-the-way references and apt illusions is impossible to summarize. In his Preface the author tells how he had been fascinated by the fact that shapes and marks can signify and suggest things beyond themselves and how he arrived at the conclusion that 'no artist can paint what he sees', discarding all conventions. To confirm his conclusions he examined his theory of perception upon which they were based. *Art and Illusion* is the record of his study. 'Why is it that different ages and different nations have represented the visible world in such ways?' Is everything concerned with art entirely subjective and are there objective standards in such matters? Representation, he asserts, is never a replica. The forms of art, ancient and modern, are not duplications of what the artist had in mind any more than they are duplications of what he sees in the outer world. The naturalistic image is a very closely knit configuration of relationships which cannot be varied beyond the range of available 'equivalences'. All artistic discoveries are discoveries not of likenesses but of equivalences which enable us to see reality in terms of an image and an image in terms of reality. Gombrich's thesis is cogently argued. A certain ambiguity, however, clings to his use of the word 'image'. In part knowledge, in part visual perception, in part unconscious projection, mental images, the author admits 'will always be an elusive welter of fleeting events which can never be communicated in full'. Hence, presumably, the eternal enigma of art and the limitations of the method of objective analysis which Prof. Gombrich himself wishes to offer in opposition to subjectivist theory.

Gombrich gives a final blow to the doctrine that the history of art is equivalent to the history of representational skill. He also disposes of Ruskin's fiction of the 'innocent eye', criticizes Thouless's right to speak of 'regression to the real object', and will have nothing to do with the 'so-called eidetic faculty' in relation to art.

T. ELDER DICKSON

The Anatomy of Judgment: an Investigation into the Processes of Perception and Reasoning. By M. L. JOHNSON ABERCROMBIE. London: Hutchinson. 1960. Pp. 156. 25s.

The perceptual demonstrations of the late Adelbert Ames, Jr., present a peculiar problem. From time to time one has to ask oneself 'just what do they demonstrate?'. Distortion rooms, rotating trapezoidal windows and the like by their very impressiveness readily evoke the response: 'so what?'. The official interpretation of Ames, Kilpatrick, Ittelson and others is that they help us to understand the influence on perception of assumptions deriving from past experience. This book is very much concerned with these demonstrations, and explores one of their implications. The author finds affinities between such concepts as 'schemata' (Bartlett), 'assumptions'

(Ames), 'hypotheses' (Bruner), 'brain models' (J. Z. Young), and 'cell assemblies' (Hebb). Despite their differences of terminology the various theorists concerned are all stressing the effects of expectancy, preparedness and learning upon perception. Mrs Johnson Abercrombie herself is interested in the possible implications of the Ames 'assumptions' for education.

She is not concerned with education in general, but attempts something much more modest and specific. After an account of the Ames demonstrations on the one hand, and the technique of group discussion on the other, she relates both to the problem of training medical students and practitioners. What the psychologist calls 'perception', and what the medical man calls 'diagnosis' have something in common. The author is interested in how medical evidence is evaluated, and how the assumptions underlying perception can give rise to errors of judgement. She is particularly concerned with the perception of radiographs, whose resemblance to Rorschach ink-blots is, incidentally, noted.

In the experiments reported, medical students were asked to comment on radiographs first individually, and subsequently in group discussions. The aim of these discussions was to teach the students to perceive detail, and to gear their judgements to the evidence actually presented on the X-ray plate. In the process the student would learn that many of the apparent 'facts' on which his individual judgements had been based proved, on examination, to be inferences. He would come to develop a more sophisticated attitude towards words like 'normal', which the group discussions showed carried an important and misleading load of ambiguity. In general, as a result of the discussions, the place of assumption—of the kind to which Ames draws attention—as a determinant of perception would become more apparent to him. While other evidence is collated by the author, of particular interest are these experiments of her own. They suggest some refreshingly new points of contact between psychology and both medical practice and medical education. The book as a whole is likely to interest all who are concerned with the relations between seeing and thinking, the process of classifying, and with the ways in which evidence is evaluated.

This admirable, well-written, and wholly unpretentious book certainly merits the attention of psychologists. It is likely to stimulate thought about the process of educating students. It may also be recommended to the general physician and to specialists in branches of medicine which do not usually have much contact with psychological theory and research. PETER MCKELLAR

Psychological Appraisal of Children with Cerebral Defects. By EDITH MEYER TAYLOR. Harvard: University Press; London: University Press. 1959. Pp. 499. 68s.

This book is the fruit of twenty-five years' experience in the psychological study of children. In the introduction the author presents her credentials: graduation from Leipzig under Felix Krueger in 1933, five years with Piaget and André Rey in Geneva, a year with Gesell at New Haven, some years as consultant to child-caring agencies in Connecticut, succession in 1943 to Dr Elizabeth Lord's post in the neurological division of the Children's Hospital in Boston under Dr Bronson Crothers, and more recently appointment as psychologist in the Cerebral Palsy unit and Research Associate in Pediatrics at Harvard Medical School.

Her attitude to her work is essentially clinical and individual. She favours 'controlled flexibility... with the emphasis on descriptive and qualifying findings'. She believes in choosing appropriate test procedures to fit the clinical situation, rather than the application of a standardized battery, and in further tailoring them, if it seems necessary, to meet individual needs.

In addition to the general description of background, aims, and methods, in the Introduction, one finds, rather to one's surprise, a brief account of the psychological findings in the follow-up study, started in 1951 by Dr Crothers, of children suffering from cerebral palsy. It seems incongruous that the assessments are reduced to ratings on a five-point scale of 'intellectual competence'. The earlier ratings were inferred by the medical investigator from the psychologist's reports, while the later were generally based on the W.B.I.S., the W.I.S.C., or the Revised Stanford-Binet. The author describes these reported results of the psychological studies as 'rudimentary', but their inclusion in this form rather reflects upon the utility of the methods of assessment advocated earlier for general scientific purposes.

Part One consists of 'Case Portraits', of which first four, various 'cerebral palsy' syndromes, are generalized descriptions based on an unspecified number of cases, but are identified by personal names, as are the last three individual cases, a meningitis, an encephalitis, and a head injury.

In Part Two techniques of examination are discussed. These are classified as: Blocks (11 pages), Pictures (14 pages), Paper and Pencil (17 pages), Fitting and Assembling (19 pages), Following, Finding and Retrieving (7 pages), Matching, Sorting, Grouping (19 pages), Designs (20 pages), Learning and Memory (20 pages), Judgement and Reasoning (25 pages), Everyday Sense, Common Knowledge, Comprehension (6 pages), Automatisms, Academic Tools, Factual Information (12 pages), Vocabulary (4 pages).

It will be observed that the usual projective techniques are not employed. The author describes the Rorschach method as 'too time-consuming and often disturbing to the patients' adding that it 'often can furnish meagre findings only'.

A preference for investigating visual perceptual and manipulative types of function, rather than language, is evident and quite explicit. The author suggests that it is unwise to probe and undesirable to treat dysphasic disorders after cerebral injury in children; she exemplifies this in the case of 'Jack', who, after head injury seems, from the case description, to have had quite clear dysphasic signs, which are not, apparently, considered sufficiently important for reference in the final summing-up.

The techniques described include a number of personally devised tests, or variants of items from standardized scales. These are generally located on a mental age scale by reference to items in the standardized scales with which they appear analogous, or by clinical impression. For the 'Visual-verbal learning test' (p. 428) one is referred for 'tentative judgements' to norms which turn out to be those for a rote word learning test, established by Rey with a 'potentially equivalent list of French words'. Of this she says, as of her other such tests, that 'it would be desirable to establish precise norms for this test, and also to study the range of individual differences which seems to prevail'.

The book is well worth reading as an exposition of the author's methods of investigation, which seem designed to be sensitive to selected facets of the effects of brain lesions upon individual children, rather than to extract broad generalisations. One would, however, have some reservations in recommending it to persons beginning work in this field. Such methods are not readily communicable in book form.

J. A. L. NAUGHTON

Epidemiology of Mental Disorder. Ed. by BENJAMIN PASAMANICK. Washington, D.C.: American Association for the Advancement of Science; London: Bailey and Swinfen. 1959. Pp. 295. 59s.

This book contains twelve papers from a Symposium organized to commemorate the centenary of the birth of Kraepelin. The first of these is, indeed, the Emil Kraepelin Memorial Lecture by Eugen Kahn. Of the many in the great tradition of French and German psychiatry around the turn of the century who were deeply concerned with the classification of mental illness, Kraepelin was probably the most persistent and successful. It is surprising to learn, of one who has always been admired for vivid presentation of case material, that he was an indifferent lecturer. He thought it was a misfortune to be a good speaker, since oratory led from facts to speculation and twaddle. This should be a comfort to the inarticulate and a warning to the inveterate congress-man. The quality of the spate of books from America based on Symposia has nevertheless justified this method of production. This particular book, though not outstanding, adds to that justification.

Since each paper is followed by discussion there is little of a detailed nature left for a reviewer to add. The paper by Clausen and Kohn on 'Relation of Schizophrenia to the Social Structure of a small city' is of particular interest in that the results conflict with those of previous studies. They found no apparent relationship between socio-economic levels of areas of a small city nor of major occupational groups and rates of hospitalization for schizophrenia. This is in marked contrast to Clark's Chicago study and to the Registrar-General's figures for Britain. They found no evidence of upward or downward social mobility in the period prior to first admission. They believe that there may be a direct relationship between size of city and the degree to which rates of hospitalization for schizophrenia correlate with such indices of socio-economic status as area of residence and occupational level. The evidence does not, as is suggested in the discussion, contradict Clark's findings. It limits their generality by introducing a variable not previously known to be relevant.

The paper by Zubin, Burdock, Sutton and Cheek on 'Epidemiological aspects of prognosis in Mental Illness' gives an admirable outline of methods for studying prognosis. Of interest too are some of the technical innovations, such as the immobility index, agreement scores and distance scores. No one can doubt the immense importance of this area of work since some of the specific therapies have been shown to have a short term, but not a long term, advantage over non-specific therapies or no treatment at all.

Other topics include personality change during the second decade in the lives of young people; a survey technique for estimating the prevalence of psychoneurosis in communities; housing environment and mental health.

The book is beautifully produced, but with a number of spelling errors not attributable to the American language and with factual errors, such as 'Woodward 1937' for 'Woodworth 1917' and the omission of the speaker's name (Macmillan) on p. 226. It is expensive for a collection of studies that are, in the main, exploratory rather than definitive.

G. A. FOULDS

Voice and Articulation. By C. VAN RIPER and J. V. IRWIN. London: Pitman. 1959
Pp. xvii + 566. 55s.

The authors of this book are two of the foremost authorities on speech pathology in the United States. They have assembled and evaluated a great deal of information on articulation and voice, sifting opinion from fact in a field where relatively little has been firmly established.

The first chapters deal with abnormal articulation, its causes, testing, symptomatology and therapy. The organic and functional disorders of voice are next described together with methods of treatment. Van Riper and Irwin consider that speech is learned behaviour monitored through automatic control, and treatment is discussed in terms of feedback theory, the patient learning to match his speech to a correct pattern which is eventually fixated. Finally there are clear and well illustrated chapters on the anatomy of the speech mechanism, phonetics and acoustics.

The book is comprehensive, of a high standard and easy to read. Its sequence would have been improved if the chapters on anatomy and phonetics had been at the beginning. The authors state that their approach to treatment is based on modern learning theory, but they give no examples of its application which would satisfy the psychologist's conception of learning theory.

This is an excellent text-book, essentially for speech therapists. It is recommended to those psychologists, particularly in the educational field, who come in contact with patients with speech disorders and to those wishing to know how the speech therapist treats these particular problems. The 868 references are valuable for research purposes; most of them are American because little research in speech pathology has been done in this country.

J. WILKS

The Behn-Rorschach Test. By HANS ZULLIGER. Bern: Huber. 1956. Pp. 200.
Swiss Fr. 25.

This is an English text of Zulliger's *Einführung in den Behn-Rorschach-Test* which appeared in German in 1941; although published in 1956, the English version seems only just to have been made available. Dr Hans Behn-Eschenburg was Rorschach's assistant at the Herisau Mental Hospital and the series of plates which bear his name was developed in close collaboration with Rorschach. The need for a parallel series was at the time felt to be more imperative than later developments have suggested. In addition, 'specialization' of the various potential series was envisaged. This, too, represents an aim which has not been realised, except that in the present book the 'Bero' (i.e. Behn-Rorschach) is described as more suitable for the testing of younger children.

Although much stress is laid on the fact of standardization against the original Rorschach plates, no details of such standardization are given. Instead individual protocols showing much or little discrepancy as compared with Rorschach results from the same person are reproduced. Since extraneous circumstances—including, of course, the very fact of re-testing—are bound to be operative, it is not surprising that the results can, at best, be regarded as comparable rather than mutually confirmatory. The two series must therefore be regarded as alternative instruments rather than as true 'parallel series' in the psychometric sense. It may well be that Bero is potentially more informative than Rorschach in specified situations, but it seems unlikely

that an experienced Rorschach user would feel it necessary to expand his scope by acquiring expertise in Bero. In particular, the author makes use of some scoring categories which are different from those of Rorschach himself or of any recognized 'school'. Chief among these is *m* for human movement seen in a small detail, for which Zulliger claims that 'one notices easily that [such responses] must originate in different parts of the subjects' mind from *M*'.

The book is beautifully produced and printed, in accordance with Huber's customary standards. There are, however, many misprints and minor errors, ranging in seriousness from failure to substitute an English for a German scoring symbol, to confusion between *F* + % and *F* %. No text. It has a strong—and rather attractive—foreign flavour. Much the same is true of the social and cultural background of the case-studies which constitute probably the most valuable feature of the book. Little, if anything, seems to have been published in English-speaking countries on the use of the Behn-Rorschach; perhaps the appearance of the English text will stimulate activity.

B. SEMEONOFF

The Ineffective Soldier: Lessons for Management and the Nation. By ELI GINSBERG, J. K. ANDERSON, SOL W. GINSBURG, JOHN L. HERMA and others. New York: Columbia University Press; London: Oxford University Press. 1959. Vol. I, Pp. xxi + 225; Vol. II, pp. xviii + 224; Vol. III, pp. xx + 340. 48s. each volume.

This book consists of three volumes of a total of 849 pages. It is an account of some of the nine years work undertaken by the Conservation of Human Resources Project, which was initiated by General Eisenhower, when he held office at Columbia University, before the Cold War called him back to Europe.

Eisenhower had been made painfully aware during the War, especially at the time of the Battle of the Bulge, that America's manpower resources were limited. (No doubt when the Cold War set in he felt this even more acutely.) His post at Columbia, therefore, presented him with an opportunity to set in motion an examination of these resources as they existed during the War. He persuaded a number of firms and trusts to sponsor the project, which was carried out by a team of ten social scientists, two consultants, and a number of assistants. Services' records departments provided much of the data which they extracted from the punched cards which they held, and additional information was obtained from questionnaires designed by the team.

The first volume, *The Lost Divisions*, contains 46 tables, 6 diagrams and two maps. It is a critical analysis of the screening procedures used at the time of enlistment, and an examination of the factors which led to men being separated from the Army at a later stage. The emphasis is placed on psychological and educational defects rather than on the physical ones.

The far-reaching effects of policies dependent upon a totally inadequate knowledge of manpower potential come out very clearly. The muddle which ensued was made worse by the hectic rate of the Army's expansion: 109 % p.a. during the war. An insistence on passing each man through the psychiatric screen meant that the time devoted to him was approximately three minutes. This first volume is a mine of information, and is a most valuable reference book, if not one to be grasped entire on one reading.

The second volume, *Breakdown and Recovery*, on the other hand, is easy reading, but of much less value. It consists of a number of life histories illustrating the various factors which led to separation. The case histories are selected, and do not pretend to be a random sample of the total population. No attempt is made to test the validities of the assumptions which are made. Thus, although the arguments are plausible and interesting, they suggest that the authors came unstuck in their efforts to provide a detailed analysis of factors leading to breakdown, and fell back on the case-history method. This volume is such that it leads one to suppose that the psychotherapist on the team had most influence in its compilation.

The final volume, *Patterns of Performance*, reverts to the design of the first one, and contains close upon 200 tables. Situational and background factors leading to breakdown are examined, but the categories of experience and criteria of performance are very broad, leaving a need for the provision of something more detailed than this, and yet of greater validity than the material presented in the second volume. Many of the tables would be improved by the addition of information about the population risk. Officers are not considered at all.

The book claims to provide 'Lessons for Management and the Nation', but one fears that management might find the search for the lesson beyond them, in spite of a chapter headed, 'The Management of Men', which is all quite true, no doubt, but does not follow from the evidence presented.

The Ineffective Soldier is an excellent reference book for all persons dealing with national manpower problems, but one wonders how many will read it.

DOUGLAS IRVINE

Health and Happiness in Old Age. By CELIA WESTROPP and MOYRA WILLIAMS.

London: Methuen. 1960. Pp. 133. 8s. 6d.

This book by a geriatrician and a clinical psychologist discusses 'the mental and physical problems of old age in a simple, popular, but scientific manner'. There are four main sections dealing with the physical and psychological aspects of: (i) the changes that occur with age, (ii) general care of the elderly, (iii) common ailments of old age, (iv) preparing for old age.

With an admirable conciseness, the authors have written what is in fact, a comprehensive and informative account of old age. It is an efficient summary, at a non-specialist level, of modern geriatrics and gerontology. As such, it will be indispensable for all those who are concerned with looking after the elderly. One would like to see a similarly comprehensive book available, at the appropriate level, for the psychologist, psychiatrist, and other specialists in this field.

J. E. ORME

Psychology for Better Living. By LYLE TUSSING. New York: Wiley; London: Chapman and Hall. 1959. Pp. vii + 496. 48s.

One should not be put off by the title of this book. In spite of some rather formidable moralizing on the back cover, even a cursory examination reveals that *Psychology for Better Living* is a serious and soundly based attempt—quoting the author's own words—'to help to foster the building of appropriate behaviour patterns that will promote social compatibility'. The systematic nature of this attempt, and at the same time its scope, may best be indicated by a further quotation from the Foreword:

'Like the present trend of psychological investigation, this text-book deals less with physiological psychology, receptive and perceptual processes, and response processes than with the other nine categories established by *Psychological Abstracts*. Most of the emphasis is on areas designated as (1) Social Psychology, (2) Clinical Psychology, Guidance, Counseling, and (3) Behavior Deviations. It is my feeling that the student is exposed mainly to newspaper, magazine, and non-technical information in these areas and should have more suitable, substantial material available.'

Incidentally to his purpose Dr Tussing has treated rather fully many topics not falling strictly within the field of psychology, such as the physical facts of sex, economic factors in marriage, the nature of social institutions, and the like. This would seem to be a good thing, inasmuch as it helps the reader to realize that psychology is part of life—that it is, in fact, concerned with people, and not with processes. The book is designed for first-year American college students, and while one feels it could not perhaps entirely replace the more systematic elementary texts, this reviewer, at least, is confident that only the most blasé first-year student in this country would fail to find it both helpful and stimulating.

Mention must be made, in closing, of an odd creature, a sort of animated sausage, who appears on almost every other page, and whose creator would seem to stand somewhere between Thurber and the artist of the Blacky pictures. At times he is endearing, and at other times repulsive—but so, after all, are most of the people one knows.

B. SEMEONOFF

Quantitative Methods in Psychology. By D. LEWIS. New York and London: McGraw Hill. 1960. Pp. viii + 558. 75s.

This substantial volume is intended by the author to provide students of (experimental) psychology with the necessary theoretical background and techniques for the quantitative treatment of psychological data. The material presented is the outcome of the author's long

experience and has been used in various forms before arriving at the particular selection and arrangement represented here. The following major topics indicate the scope of the book: Variables—Linear functional relationships—Logarithms and logarithmic scales—Non-linear functional relationships (referred to by the author as 'complex functions')—Differentiation—Integration—The normal curve—Binomial and Poisson distributions—The χ^2 , t and F distribution functions—The 'goodness of fit' of curves to experimental data—The application of equations.

Many of these topics are treated in considerable detail and where possible appropriate psychological examples are employed. Non-linear growth curves are illustrated by a general form proposed for such curves by Gompertz in 1825. Differentiation is employed to derive the formulae for the best fitting equation to represent empirical data. Space is given to an examination of the properties of the Gamma function or generalized factorial (i.e. $n!$ in which n is not an integer). Examples of Analysis of Variance include the partitioning of variance in curve fitting and the use of variance techniques to test the parallelism of a set of lines. Examples of the application of equations to problems of the Behavioural Sciences include an extended treatment of the application of Fechner's law and its contemporary derivatives, the theoretical analysis of the photo-receptor processes in the eye (after Hecht), and a compact survey of 'Hull's Theoretical System for Learning'.

It will be evident that in many respects Lewis has provided a unique collection of information. The problems selected for treatment are dealt with effectively in a stimulating manner and an excellent index helps to co-ordinate information discussed in different parts of the book. The general emphasis is placed upon problems and techniques well over towards the domain of applied mathematics and the reader who is not already moderately well-versed in this subject will sometimes find that too much background knowledge is assumed. For instance Jacobian transformations are employed at several points in the book, although it has to be admitted that intelligent use of the index would allow the reader to grasp the purpose and meaning of this particular technical device. The author indicates in his introduction that this book is not considered to be a substitute for any regular text on statistics or mathematics.

To this reviewer the important problems posed by quantitative methods in psychology are by no means covered by the techniques of applied mathematics. Qualitative and logical problems which do not fit neatly into the scheme of applied mathematics receive little attention here. A more serious limitation lies in the fact that there is no reference to non-parametric techniques, in general. Yet it may be argued that for some time to come much of the raw material for quantitative psychology will be of the type which will provide increasing scope for these methods with proportionately less emphasis on what might be referred to as the classical methods of ideal quantitative systems.

ROLAND HARPER

Statistical Analysis in Psychology and Education. By G. A. FERGUSON. New York and London: McGraw Hill. 1959. Pp. vii + 347. 54s. 6d.

Prof. Ferguson has produced a clear, workmanlike and up-to-date introduction to statistical analysis for psychologists and educationalists.

Although his emphasis is on concepts, he none the less provides working instances at all points and rounds off each chapter with a set of useful exercises. He is also successful in the difficult compromise, necessary for the help of readers who are not mathematicians, between merely stating mathematical formulae and deriving them. These derivations are always very simply stated and clear to follow. This is particularly noticeable in his discussion of rank correlation methods. The book as a whole provides a stimulating contemporary point of view.

As a reader of numerous higher degree theses, I am fully in sympathy with his injunction that experiments must be designed to enable the worker and reader of the thesis to draw inferences. How many times do we find the statement in theses that the results do not permit of 'generalizations from the particular sample of individuals who participated in the study'! It is such pieces of advice which make the book of great use to the research worker at, say, the master's level.

Turning to one or two specific topics. I like Prof. Ferguson's introductory treatment of the nature of probability and his similarly concise discussion of the essential ideas of sampling. Here his treatment of biased and unbiased estimates and degrees of freedom would seem to be very helpful to the beginner.

Five useful chapters are produced on correlations and attention given to rank correlation methods seems particularly helpful. Here Prof. Ferguson describes among other coefficients, the coefficients of concordance and of consistence useful in connexion with paired comparisons. In chapter 13 he carries on to consider contingency tables and point correlation methods. There is also a short chapter on non-parametric tests. This falls naturally into place after his introduction early in the book to the different kinds of statistical variables.

Although he provides finally a chapter on partial and multiple correlation, he does not take up factorial analysis. This perhaps is one of the few criticisms one can make of the book since in fact so many first researches carried out at the bachelor's and master's level involve the use of factorial methods. A simple algebraic (as opposed to statistical) treatment and working account say, of the centroid method, together with an account of the tests of significance of the residuals might have helped, but no doubt Prof. Ferguson would prefer to leave this to the existing books on this topic.

E. A. PEEL

Human Factors: The Journal of the Human Factors Society of America. Edited by STANLEY LIPPERT. New York and London: Pergamon Press. Vol. 1, no. 1, September 1958. Subscription price \$10. per volume.

The application of the biological sciences to the study of human performance at work and of machines and equipment for work came to the fore as a result of wartime problems created by equipment for the armed services. It has since, however, been steadily growing in industry under the names 'human engineering' and, more recently, 'ergonomics'. It represents the attempt by anatomists, physiologists and experimental psychologists to consider practical problems in the light of fundamental knowledge and to learn in their turn from doing so. It is thus more than applied anatomy, physiology or psychology in the sense that not only the methods but also the principles of the disciplines concerned are brought to bear, and this in what is essentially an atmosphere of enquiry and research.

Although much of this work has been, and will continue to be, published through the media of the several disciplines concerned, journals specially devoted to it are beginning to appear. The British-published, international journal *Ergonomics* (see this *Journal*, 49, p. 175) has now been followed by an American counterpart. The journal *Human Factors* is the official organ of the Human Factors Society of America and is a very welcome addition to the literature. In the first number, after a short introduction by Prof. L. E. Morehouse, an outline of trends in the area and a plea for the maintenance of scientific standards are given in a leading article by J. M. Christensen. There follow two special articles, one, by A. E. Hickey and W. C. Blair, discussing a cybernetic model of monitoring tasks, the other, by E. B. McFadden and J. J. Swearingen, giving data upon the forces that can be exerted on lever-type door handles as fitted in aircraft. After these is a short report by J. A. Kraft of a survey of 'human factors' studies in the American aircraft and missile industries, and a valuable annotated bibliography by Patricia S. Allen and E. V. Saul from that acknowledged centre of bibliography in this area Tufts University, Medford, Mass. The number ends with a book review, two letters to the editor and a brief editorial.

The journal has wisely started fairly small—48 pages in the first number—and has managed to combine an obvious seriousness of purpose with a lightness of touch. If over a succession of numbers it is representative of current research in the area it will undoubtedly contain many articles that psychologists ought to read, even if they are not primarily interested in problems of engineering psychology.

A. T. WELFORD

Journal of Child Psychology and Psychiatry and Allied Disciplines. London: and New York: Pergamon Press. Vol. 1, no. 1, January 1960. Subscription price to private individuals £5. 5s. per annum.

This first issue of a new journal, which is the official organ of the Association of Child Psychology and Psychiatry, is edited by Drs C. B. Hindley, Elizabeth Irvine, and Emanuel Miller. The aim is to 'bring together original papers concerned with child, from such diverse disciplines as psychiatry, psychology, paediatrics, psychoanalysis, social casework and sociology'. It is hoped

to make the journal international both in readership and authorship, and the editorial and advisory boards are made up of well-known people from all over the world.

The first article is a 'Discourse on Method in Child Psychiatry' by Emanuel Miller, one of the editors, which consists of his address from the Chair to the Association in January 1959. Prof. Leo Kanner follows with a salutary article suggesting that behaviour symptoms do not always indicate psychopathology. Drs A. D. B. and A. M. Clarke usefully summarize recent work on 'early deprivation' in children, which is followed by a report of another investigation in the same field from the Department of Child Study in Birmingham. This report concludes that early physical separation and prolonged institutionalization do not necessarily lead to emotional difficulties or character defects, but are likely to if, in addition, the child remains rejected and unwanted.

One is particularly attracted by the mixture of theoretical, experimental and clinical articles in this issue, which has a uniformly high standard of presentation, readability and interest. If this is maintained, yet another journal will have to be added to the growing list of essential reading.

It is to be hoped that the publishers can contrive a more durable way of fixing the cover, and that more time be spent on proof-reading.

RALPH HETHERINGTON

Grundlagenstudien aus Kybernetik und Geisteswissenschaft. No. 1, January, 1960.
Pp. 32. Stuttgart: Privately published. D.M. 2.00.

The intending reader of this new journal might well wonder about the juxtaposition in its title of cybernetics and 'Geisteswissenschaft'. However, an introductory leaflet explains that this is intended to draw attention to the role of cybernetics as a bridge between the methods of the natural sciences and the humanities. The Editorial Board lay stress on the inter-disciplinary interest of all the present and expected contributions.

There are four articles in the first number of the journal. The first two, being concerned with aesthetics and some special computer programming problems, are of only peripheral interest to psychology. The other two articles are directly concerned with psychology. In one of them F. von Cube describes the application of Shannon's formulae to studies of perception and learning. He bases his survey to a considerable extent on papers which appeared in *The Journal of Experimental Psychology* in the years 1951-56 and which will be familiar to all interested in information theory. In the other article H. Frank sets out to describe 'informational psychology'. His is the only paper with French and English summaries, although one would wish that the English summary were a little more helpful (e.g. the first two sentences run: 'How long a perceived sign is present in the conscience may be fixed rather exactly to 10 sec. The time necessary for the perception of a sign of subjective probability p /probability for an individual! is found by combination of several processes.')

Those concerned with cybernetics and information theory will watch with interest the development of this new continental journal. It remains to be seen to what extent its ambitious aims can be fulfilled.

W. SLUCKIN

THE ACCURACY OF 'REPRODUCED' CONVERSATION*

By T. H. PEAR

Emeritus Professor of Psychology, University of Manchester

In any witness's account from memory of a heard conversation, important factors are his motives and his mental apparatus of recall and reproduction. Since rendering the gist of a conversation involves editing by omission, all the factors in Freud's 'dream-work' may have combined to distort the account given in a court of law, especially by a witness who extensively uses 'direct oration'. The numerous experiments on the reliability of testimony for seen events contrast with the few in which heard events have been the material. Is auditory imagery as reliable as visual? A social psychologist's view is given of the treatment of conversation by historians, biographers and novelists. The possibility that a strict training in scientific psychological method may impede the writing of an imaginative story, and the use of tape-recordings are discussed, along with the ethical and legal issues arising from the use of tape-recordings as evidence.

I. INTRODUCTION

When, about 1906, I began to study simultaneously physics and psychology, subjects in which, at that date, 'accuracy' may have had rather different meanings, my attention was attracted by a pattern of problems which still seem as topical, socially important and neglected as in those days. An occasion which caused me to formulate some doubts, still unresolved, was a Press report of a criminal trial. Allowing for tendencies to distort memories in ways which Freud was at that time describing to a handful of Austrian doctors, let me indicate the focal point of these challenging puzzles.

A servant girl had asserted under oath that a long time ago—months, I think—A had said this and B replied that. Nobody seemed to doubt her ability to reproduce A's and B's exact words. At that time, I formulated no questions to myself in psychological terms: had I done so, one would have been 'Did the legal experts concerned with the case believe her to be reporting the words which she and others had spoken?' A country dweller myself, I knew that few simple people ever tell a story involving themselves in a form other than *oratio directa*. If this direct oration is subsequently recalled, reconstructed, perhaps with visual images of the speakers' facial expressions and gestures, and verbalized hastily and excitedly, are there not psychological justifications for doubting if the witness is speaking the truth and nothing but the truth?

Even if we ask Pilate's question, we cannot honestly evade the problems. Not only the witness's motives, but his mental apparatus of recall and so-called reproduction are all-important. A 'visualizer' myself, even today I would readily attest that one of my schoolmasters had a short pointed beard and carefully tended hands: I can 'see' them now. I doubt if I could repeat accurately, for my auditory images are as infrequent, faint and sketchy as my visual images are importunate and detailed, any sentence he spoke during the six years that he taught me—and his remarks were often characteristic of him. Might an 'audile' have succeeded where I fail?

* Extended version of a paper read at the Cambridge Conference of the British Psychological Society, 4 April 1959.

II. THE VALIDITY OF PSYCHOANALYSTS' EVIDENCE

When psychologists trained in laboratories became acquainted with Freud's early teachings, they naturally asked questions about the validity of some psychoanalysts' evidence. It was based on conversation—usually hours of it—but for obvious reasons, verbatim accounts in shorthand had not been taken, nor in those days was a tape-recorder available. Psychoanalysts relied on their memories, aided perhaps with notes, to assay the patient's ramblings, but how often did they turn Freud's penetrating searchlights on their own minds; their motives and the workings of their memory apparatus? How many considered the effects on their 'reproduction' (so-called) of a patient's revelations, of symbolism, condensation, dramatization, and especially of regard for presentability? If the production of 'gist' involves editing by omission, is the comment unfair that, given a tape-recording of only one hour's free association, a Marxist or a non-medically-oriented social psychologist could distil from the 9000 words economic and 'class' as well as sexual implications?

III. THE CREDIBILITY OF REPRODUCED CONVERSATION

Unless I have misunderstood them, the few lawyers with whom I have discussed the psychological bearings on court procedure of these doubts believed that the average untrained person could be expected to produce the gist of a remembered conversation. Yet certain court proceedings, reported with reasonable detail in the more responsible of our newspapers, suggest that a barrister with scruples like mine might inject doubts into the minds of a jury. Here, for what they are worth, are my memories of accounts of court cases:

When a well-known ex-M.P. was being examined in court, part of the evidence, subsequently read out as news on the radio, was in direct oration. In another case, two men, *A* and *B*, had discussed an important matter, heatedly expressing their differences. After *A* had left, *B*, according to his account, had caused a typed 'record' of the conversation to be made while it was fresh in his mind, *A* was not asked to make one. To the comment that in such a case it is only the gist that matters, one might object that by scholars, *précis* making is regarded as an art in which few succeed brilliantly, even when the material is placed before the 'gist-maker' in cold print, and he is neither cross-examined in public, nor made to feel socially inferior.

IV. THE INFLUENCE OF VISUAL AND AUDITORY IMAGES

Numerous experiments have demonstrated the unreliability of testimony of seen events. Some visualizers have teased out from their images of things once seen, the effects of those processes of 'dream work' which, in their political cartoons, Low and Vicky use consciously or unconsciously. Yet some aspects of persons whom I saw long ago resist the usual memory-'condensation' of images which in one or more respects resemble each other. I am puzzled by the fact that, though I am not an anatomist, at first sight of persons who interest me, I note, visually but not verbally, peculiarities of their hands and teeth. This habit owes nothing to systematic training in observation of such details. Since I would be willing to swear to the existence of these, not easily altered without warning and some knowledge of the techniques of disguise, my evidence might be important in a court trial.

In contrast, my memories of speech are not formed in this way, though I have occasionally suspected the fusion in memory of similar voices. After 25 years' casual observation, I usually notice the timbre of a voice, especially if it seems beautiful or ugly, and may remember it in borrowed visual terms, of colour, shape or both, I perceive intonation, vocabulary, pace, and 'accent' suggesting region or social class, but seldom remember the actual words spoken, unless they indicate nation, region, vocation or class, or are annoyingly or amusingly incorrect. Since accurate memory of words and phrases is so difficult for me, I might perhaps be tempted to suspect, like the fox who lost his tail, all so-called reproductions in direct oration, yet when someone whom I knew well recalled what her tutor said to her, and what she replied, at least 40 years ago, I wondered if her auditory or auditomotor images were at times as reliable as my visual images. Might not auditory images in some persons also resist time's ravages? On this point I recall no experimental evidence: now that recordings of speech can be uncannily faithful, this field of research seems attractive.

V. HISTORIANS, BIOGRAPHERS AND NOVELISTS

The social psychologist's attitude to the treatment of conversations by students of history and biography, and by novelists, is too seldom made clear. Who believes in the factual accuracy of records of conversations with the famous dead? Some people now smile at the 'Scrap of Paper' and the 'Contemptible Little Army' but in World War I most British people believed that the literal German equivalent of these words had been spoken by highly placed enemies, and our Government did not strive to correct them. If there were a French equivalent of *1066 and All That*, 'Let them eat cakes' might appear in it. One writer asserts that Marie Antoinette meant 'brioches', and André Castelot writes that according to Mme de Boigne, Madame Victoire, daughter of Louis XV, said 'with tears in her eyes, during a period of scarcity, when someone was talking of the sufferings of the unfortunate people who lacked bread: "But if they could only resign themselves to eating piecrust!"'

(In quoting these alternative renderings, I do not suggest the absence of others equally amusing and regarded as more credible.) Ought psychologists to be sceptical about 'literary' reports of conversation? Sir Harold Nicolson wrote:

It is delightful to overhear the conversation at Agathon's supper party, or to listen to the chatter in the wrestling school of Taureas, or to hear Dr Johnson thundering away in Fleet Street... or Anatole France in carpet slippers. These reporters of conversation need not be as gifted as Plato; in fact, it is preferable for them to be unpretentious or even naïve, and from time to time they should record remarks the true significance of which they have obviously failed to understand. This adds conviction to their reporting.

Was Sir Harold smiling unilaterally? These later paragraphs seem serious:

Among those who have accurately and modestly recorded the conversation of their heroes, I should give top marks to Johann Eckermann (Goethe's secretary) and Philip, 5th Lord Stanhope, who recorded the Duke of Wellington's conversations for 21 years.

Both Eckermann and Stanhope are, compared to Boswell, untarnished mirrors; there is nothing oblique about their reflection, and the impression derived from such records is often more detailed and valuable than the most detailed biography.

So, inevitably, we have arrived at Boswell.* Joseph Wood Krutch has lovingly

* I thank Prof. L. C. Knights for help in the following examination of assertions that Boswell was an untrustworthy reporter.

studied a mass of writings about 'the first boswelliser the world has ever known'. From Krutch and from Boswell himself we know that he touched up his accounts, seeing to it that 'Sir' occurred in most of the reported remarks, especially of the 'knock-down' kind, and that Boswell translated into 'more stately and more decorous Johnsonese' some of the great man's remarks about sex and physiological functions. It is important to read Boswell on his own accuracy of reporting:

I was so wrapt in admiration of his (Johnson's) extraordinary colloquial talents, and so little accustomed to his peculiar mode of expression, that I found it extremely difficult to recollect and record his conversation with its genuine vigour and vivacity. In progress of time, when my mind was, as it were [and here comes the famous phrase] strongly impregnated with the Johnsonian ether, I could with more facility and exactness carry in my memory and commit to paper the exuberant variety of his wisdom and wit.

Boswell often records whether in a particular instance he did or did not write notes of a conversation soon after it occurred. 'Sometimes' he made them 'very soon after' a conversation which he considered to be significant. . . in a simple system of abbreviation omitting many syllables or words which can readily be 'supplied'. (The last sentence may make an experimental psychologist wince four times.) Boswell used this shorthand not to record words as they fell, but to save time in making, at the first opportunity, notes to be incorporated in his *Journal*.

My method is to make a memorandum every night of what I have seen during the day. By this means I have my materials always secured. Sometimes I am 3, 4, 5 days without this journalising. When I have time and spirits, I bring up this journal as well as I can in the hasty manner in which I write it. . . the interval allowed to elapse between the notes and the 'writing-up' tended to lengthen from a few days to a few weeks.

Krutch tells us that

in *Boswelliana*, Johnson is reported to have said of 'old Sheridan' and his schemes for the improvement of elocution 'He is like a man attempting to stride the English Channel' and then 'It is like setting up a candle at Whitechapel to give light at Westminster.' In the *Life*, Johnson is made to say 'Sir, it is burning a farthing candle at Dover to give light at Calais'.

A pretty example of condensation!

I have been privileged to ask some successful novelists how, and for what purposes, they invent conversation. To some of them, 'Did A (a real, living person) ever say just that?' seemed a reasonable question for a scientist, lawyer or judge, to ask, but of little interest to a novelist trying to impregnate himself with the ether of his imagined hero. Some referred respectfully to 'psychological truth', but a psychologist would use neither of these words with the meaning they seemed to bear in a novelist's mind; correspondence between the impression produced on the reader of average intelligence and sensitivity, and that which the writer tried to convey by words given to the character who spoke them. One writer said, in effect, 'The novelist builds up the character: if the reader accepts it, a confidence trick has come off'. Perhaps reviewers are more resistant than the average reader to confidence tricks. This might explain why one wrote, of the dialogue in I. Compton Burnett's novels, 'Occasionally I find the same fascination in reading this prose that comes from looking at a knitting pattern. Instead of the hard, short sentences, there might as well be "k6, p8, pssso"'. Her admirers would object that she hints at deeper, or higher, levels of personality and character to be inferred by sympathetic readers. E. M. Forster, who knew Elizabeth of the German Garden, thought her insensitive to 'the

lower forms of truth', for example, a correct description from memory, of his necktie.

One is sometimes asked why psychologists, who depend so much on conversation—this appears to be the belief of the man in the street—seldom write novels. I believe that strict training in psychological observation of oneself and others may diminish or even destroy the ability to invent conversation. A psychologist, trying to do this, seems likely before long to reflect that the heroine couldn't say the sentence he has attributed to her, and since she never existed, the whole exercise is pointless. I once heard Dr C. S. Myers, describing to a psychiatrist a convincing example to support a suggestive hypothesis, break off with 'Sorry; I've "fused" two different cases'. For such a slip Boswell might have thanked God.

I am inclined reluctantly to agree with the view that while it may always be profitable to know how a novelist's mind works, to speculate about the 'psychology' of his characters may be a waste of time. If to impersonate on the radio a famous character is another kind of confidence trick, it should interest the psychologist to observe when and why the trick succeeds, or fails—for him. A popular, intelligent and irritating broadcaster's speech bristled with so many cultivated mannerisms that it was an easy attraction for stage impersonators. Usually, however, the sentences attributed to him were so fatuous that knowing him and his profession, I could never enjoy the skit.

I remember—accurately I think—a few phrases because they seemed unusual in their context: Prof. Samuel Alexander, with one of his serious twinkles, describing a certain original thinker as a clever donkey; Sir Charles Sherrington, renowned for pauses in the middle of a sentence when he fastidiously sought the only really suitable word, criticizing an impressive procedure which claimed to be scientific as 'just... hokey-pokey'; Bertrand Russell reminding a restive, reddening Rutherford that the atom-models he was displaying were conceptual constructs, and Dr W. H. R. Rivers asserting that in high-table conversation, Cambridge dons are concerned to discover if a statement is true; Oxford dons to formulate an epigram about it. Yet were these phrases, when spoken, as striking as they seem to me today, quoted out of their original visual and auditory context? Have they resisted the forces of the 'dream-work'? I believe they have, but am sobered by thoughts of that fusion between Whitechapel and Dover....

One of our leading newspapers recently published discussions between their reporter and a number of well-known persons, including the Archbishop of Canterbury and Jean-Paul Sartre: were the actual words recorded at the time? Martin Mayer's *Madison Avenue, U.S.A.* is based on many interviews with New York advertisers and the claim is made that these accounts are correct. It seems not improbable that a tape-recorder was used: were the tapes edited? This question becomes more important every week, judging from newspaper editorials and correspondence. The use of automatic recorders raises problems of theoretical as well as practical interest to psychologists as well as to judges. One concerns the identification of voices: what happens if the speaker alleged to have spoken a sentence asserts that he was impersonated? In a recent court case, the answer required to the question 'Had an actor's voice been successfully copied?' was inevitably based on subjective criteria. Imagine the same question being asked of a layman about a painting! When a story,

true or imaginary, is broadcast by an unpractised speaker, a recording is often made and edited in a manner which has been fully described on Network Three, for the benefit of amateur tape-recorders. There is an attempt to improve the manner and yet to retain the speaker's individuality. Even if—and this seems unlikely—no important word is 'improved', annoying hesitations and tiresome repetitions may be deleted. Then, if the story sounds too smooth, 'well now's' can be added to taste.

It appears that the admissibility in courts of tape-recordings, which varies in different countries, is being actively discussed in many circles.

VI. ELECTRICAL RECORDINGS: ETHICAL AND LEGAL IMPLICATIONS

Up to now, ethical questions have been ignored, yet 'What right has anyone to tape-record a conversation without the speaker's knowledge and previous or subsequent permission?' has naturally been raised in psychological circles. The surreptitious use of both the tape-recorder and the one-way screen are regarded by some as half-way to '1984'. Dr Robert Jungk's *Brighter Than a Thousand Suns* has been severely criticized by some scientists qualified to judge his assertions about physics. To a psychologist the book is interesting, since the author distinguishes between the degree of authenticity of the reports of speech; for example, 'It is said that A said...'; 'A said, observed, remarked...'; a conversation is reported with one speaker's warning that the words printed may not have been those actually used. Tape-recordings in American 'hearings' are freely cited. The conversation of German physicists interned in England, on hearing that the atom bomb had been dropped on Japan, was reported by a British army officer who, Jungk alleges, did not understand physics well enough to give a reliable account. In one instance the German officer's conversation was tape-recorded without their knowledge.

A question which perhaps raises ethical and legal problems concerns the degree of trust which can be placed on any person's account of an important discussion in which he came off best—or, since he may be gifted with a sense of dry humour—worst.

The query, tiresome enough, 'Did he say it?' is at times less important than 'Did he mean it?'. There are noticeable individual differences in ability to interpret facial expression, gesture, and speech melody, even in a fellow countryman. But what if a foreigner speaks his own language well and yours badly? Psychiatrists, lawyers and police officers, not to mention diplomatists, may often ask themselves this question. Students of communication and social psychologists might help them to find more satisfactory answers.

In a discussion following the address on which this article was based, Prof. Mace's expressed opinion was that while scientists desire, as a basis for hypotheses, experimental data rather than anecdotes, true anecdotes may lead to focused observation and detailed experimental research. I venture to hope that this hint may be taken.

REFERENCES

- ALLPORT, G. W. & POSTMAN, L. (1947). *The Psychology of Rumor*. New York: Henry Holt.
 BLACKBURN, J. M. (1945). *Psychology and the Social Pattern*, pp. 48–51. London: Kegan Paul.
 CASTELLOT, A. (1957). *Marie Antoinette*. London: Vallentine Mitchell.
 FREUD, S. (1937). *The Interpretation of Dreams* (revised edition). London: Allen and Unwin.
Guardian Editorial, 'Subpoenaed'; 4 December 1959.

- HAWARD, L. R. C. (1959). The psychologist in a court of law. *Bull. Brit. Psychol. Soc.* **39**, 1-8.
- JUNGK, R. (1958). *Brighter Than a Thousand Suns*. London: Gollancz.
- KERR, M. & PEAR, T. H. (1931). Synaesthetic factors in judging the voice. *Brit. J. Psychol.* **23**, 167-70.
- KRUTCH, J. W. (1948). *Samuel Johnson*. London: Cassell.
- MANSON, R. H. & PEAR, T. H. (1937). The testimony of conversation. *Brit. J. Psychol.* **27**, 277-91.
- MAYER, M. (1958). *Madison Avenue, U.S.A.* New York: Harper and Brothers.
- NICOLSON, H. *Observer*, 28 December 1958.
- PEAR, T. H. (1922). *Remembering and Forgetting*. London: Methuen.
- PEAR, T. H. (1931). *Voice and Personality*. London: Chapman and Hall.
- PEAR, T. H. (1937). The place of imagery in mental processes. *Bull. John Rylands Library, Manchester*, **21**, 193-214.
- PEAR, T. H. (1939). *The Psychology of Conversation*. London: Nelson.
- PEAR, T. H. (1957). *Personality, Appearance and Speech*. London: Allen and Unwin.
- WINGFIELD-STRATFORD, E. (1936). *Good Talk*. London: Lovat Dickson.

(Manuscript received 11 January 1960)

APPARENT CHANGES IN THE VISUAL THRESHOLD: CENTRAL OR PERIPHERAL?

BY NORMAN F. DIXON

University College, London

An experiment is reported which examines the issue of perceptual defence by measuring the threshold of one eye during the subliminal presentation to the other eye of anxiety-producing stimulus material.

By comparing the discrimination thresholds for each of two complementary colours with the ratios of these colours that are required to produce an achromatic mixture, under two conditions of subliminal stimulation, an attempt is made to distinguish between peripheral and central changes as likely factors in perceptual defence phenomena. The evidence from this exercise suggests very tentatively that 'perceptual defence' depends upon central as opposed to peripheral mechanisms and that these underlie changes in the decision processes of recognition.

INTRODUCTION

In an earlier paper by the author (1958*b*) an experiment was described which investigated the controversial issue of perceptual defence by a somewhat new approach. Instead of adopting the traditional technique of comparing recognition thresholds for 'emotional' words with those for 'neutral' words, the practice was adopted of measuring the visual brightness thresholds for one eye during subliminal stimulation of the other eye, with, respectively, emotional and neutral items of stimulus material. It was claimed for this method that since the subject never consciously perceived the stimulus material and was not required to make a verbal response, the possibility of voluntary suppression was ruled out.

The results from this experiment suggested that the visual threshold was generally raised for emotional as opposed to neutral words. They also suggested, in line with findings from a research by Blum (1955), that a period of heightened sensitivity precedes that of lowered sensitivity for anxiety-producing words. Finally, the results showed a sex difference, in regard to threshold changes determined by particular words, such as might be predicted on *a priori* grounds. The apparatus used for this experiment, which involved the methods of closed loop control, has been described in a separate paper (Dixon, 1958*a*).

Subsequent to this research, Dr Haider of the University of Vienna carried out a similar experiment using the same apparatus. In Haider's experiment the threshold for one eye was measured during subliminal stimulation of the other eye with each of the four words, 'cancer', 'breast', 'stance' and 'recant'. The results showed a one-to-one correlation between the threshold level for each of these words and the order in which these same words had been ranked for emotional unpleasantness by an independent group of fifty subjects.

On the assumption that these results were not fortuitous at least four sorts of explanation suggest themselves. Either there was a change in the sensitivity of the peripheral receptor, or there was a central blocking of that channel which mediated the anxiety-producing stimulus, or there was a change in the decision process underlying the making of a response, or, finally, there was some combination of these

possible effects. It is noteworthy that these hypotheses are quite compatible with recent findings regarding the regulatory function of cortico-reticular mechanisms (Samuels, 1959), and it is in the context of this evidence that they are exemplified in the following highly simplified block diagrams.

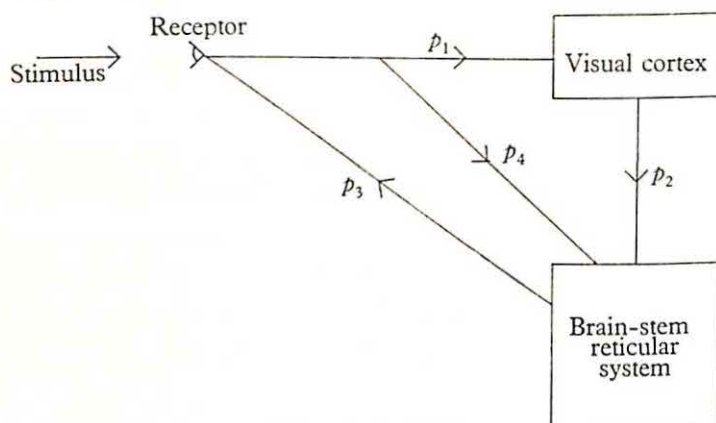


Diagram 1

According to this model cortical discrimination of sensory impulses in the specific pathway p_1 evokes, through the projection p_2 and reticular pathway p_3 , inhibition of the peripheral receptor. This inhibition will in turn reduce activity in the collateral p_4 , thus reinforcing the suppressive characteristics of the reticular system.

In support of this hypothesis Dodt (1956) and Granit (1955) have demonstrated facilitatory and inhibitory effects on retinal ganglia and optic tract by reticular stimulation.

A second possibility is depicted in Diagram 2.

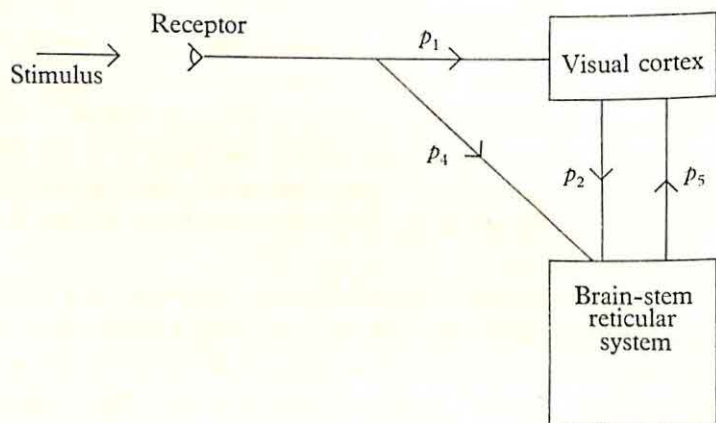


Diagram 2

Here the cortex as a result of discriminating sensory impulses in the specific pathway p_1 controls via its projection p_2 activating mechanisms of the brain stem, thus influencing its own level of arousal via the return path p_5 . There is considerable evidence from animal studies for a regulatory mechanism of this kind (French & Hernández-Peón, 1955; Gangloff & Monnier, 1956, *et al.*). An important feature of the mechanism is that the impulse velocities in the specific pathways exceed those

in the reticular formation (French, Livingston & Hernández-Peón, 1953; French & Hernández-Peón, 1955). This, as suggested by the experimental work of Ingvar & Hunter (1955), makes it possible for the occipital cortex to influence brain stem potentials initiated by direct optic collaterals, p_4 in Diagram 2. A second important feature, also supported by experimental evidence (Gellhorn, 1954; Lindsley, 1956), is that consciousness of the specific afferent supply can only occur when the sensory impulses reaching the cortex are coincident with reticular arousal of this area.

Finally, it may be added for this model that there is much evidence of a psychological kind for the existence of just such a filtering mechanism in sensory channels as these neural arrangements would provide (Broadbent, 1958; Bender, 1952; Pettigrew, Allport & Barnett, 1958).

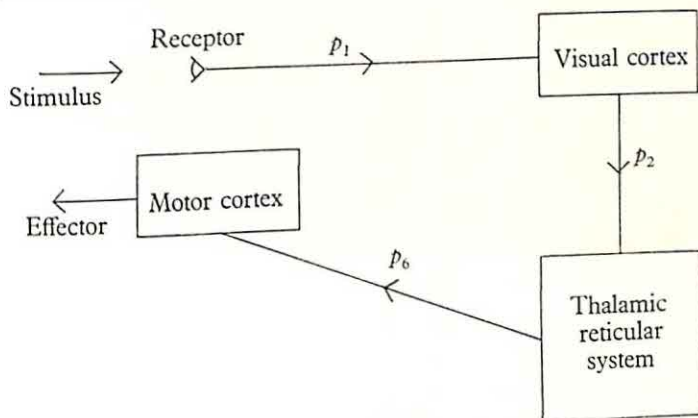


Diagram 3

This model differs from the previous ones in as much as the controlling influence of the cortex acts either directly or through the thalamic reticular system upon motor outputs. Neurophysiological evidence for this mechanism has been provided by the work of Bernhaut, Gellhorn & Rasmussen (1953).

It will be clear from the foregoing highly speculative discussion that far from being an impossibility (Howie, 1952) threshold regulation as a result of discrimination below the level of awareness is a psychological phenomenon for which several neural mechanisms exist. Any of these mechanisms working independently or in combination could, it is suggested, produce the effects reported in 'perceptual defence' experiments. Nor are these the only possible mechanisms; thus Babskii (1943) and Granger (1957) have described threshold changes ascribable to autonomic activity, while Therman (1938) has provided evidence for changes in retinal sensitivity produced by the local application of adrenaline, a substance which is also believed to produce the cortical arousal response by its action upon a specific portion of the midbrain tegmentum, (Rothballer, 1956).

The experiment reported here has a more modest aim than that of attempting to discriminate between these different mechanisms. It is merely concerned to discover whether the site of threshold change is likely to be the peripheral receptor or more centrally located.

It does this by comparing changes in threshold for two complementary colours with the change in the ratio of these colours to produce apparent achromaticity.

The rationale for this method is that if any observed change in threshold for either of the two complementaries is paralleled by a change in the ratio of these colours for achromaticity it suggests the possibility of a change in sensitivity of the peripheral receptor system. If, on the other hand, the threshold for red, say, goes up but the amount of red in the red/green achromatic mixture remains the same it would seem to exclude the possibility of a peripheral change (e.g. model I) and imply therefore a central change (e.g. models II or III).

In the context of these more general hypotheses the research examines, first, the form of threshold change over time, as a function of different sorts of stimulus material presented below threshold to the other eye, and, secondly, the effect upon differential colour sensitivity in one eye of subliminal stimulation with light of different wavelengths to the other eye.

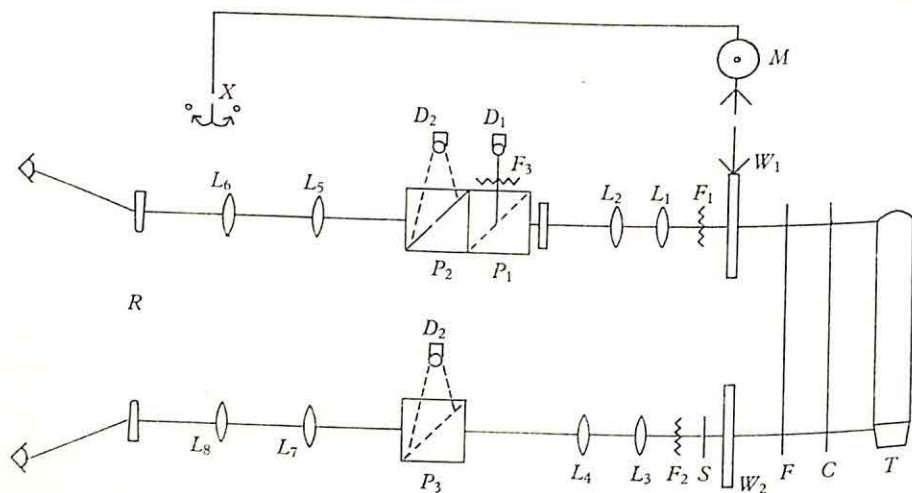


Fig. 1.

APPARATUS AND METHOD

The apparatus is shown in Fig. 1.

Light from a 1 kW. tungsten lamp T passing through a condenser C and infra-red filter F is divided between the two pairs of neutral continuous function log wedges W_1 and W_2 . That passing through W_2 produces an image on the inner face of Lummer prism P_3 of stimulus material in the film-strip holder S . The wavelength of this light is determined by an Ilford tricolour filter at F_2 . The image within the prism is focused by means of the lenses L_4 and L_3 . Light passing through W_1 and colour filter F_1 is focused by L_2 and L_1 upon the inner face of the Lummer mixing prism P_1 . Also incident upon this inner surface of P_1 is light from the tungsten lamp D_1 which passes through the Ilford Minus filter F_3 .

Between prisms P_1 and P_2 is a diffusing glass (opal-flashed). P_2 consists of a Lummer prism whose internal surface is, apart from a small circular central aperture, opaque to light entering from P_1 and reflecting of white light from the lamp D_2 .

By this arrangement the subject who regards the two prisms through the stereoscope R and focusing system L_5, L_6, L_7 and L_8 'sees' with his right eye a coloured image of the stimulus word on a lighted background and with his left eye a circular patch of light on a slightly less bright background.

The experimental method involves, first, determining the subject's threshold for stimulus material presented to the right eye, and, secondly, the controlling by the subject of the colour of the circular test patch visible to his left eye. Control is mediated by a lever X operated by the subject's right hand which, via a motor and the variable wedge W_1 , controls the intensity of

light incident upon the near face of the mixing prism P_1 . Values are so arranged that when W_1 is moved from one extreme to the other the colour of the spot changes from pink to blue green. At some intermediate position the spot is seen as achromatic or grey. This spot can be indicated by the subject via a press button on top of the control lever.

The thresholds for red and green, and the ratios of these colours required for seeing grey, are indicated on a dial, which moves consonant with the subject-controlled wedge W_1 . A light signal on the dial indicates the grey point indicated by the subject for each excursion from one colour to the other.

To obviate the subject working on anything other than visual cues in making his colour discriminations the speed of the wedge motor varies randomly from trial to trial. This is achieved by a uniselector activated by the subject's control lever which switches different resistances into the motor field circuit. Each of these resistances is itself variable by E .

The various values used for the filters, etc., were as follows:

Wedge W_1 —0.10 density units per inch with a range from 0.1 to 0.63.

Filter F_1 —Ilford tricolour no. 203 which transmits freely from 6000 Å. to infra-red and whose absorption increases with decreasing wavelength from 6100 Å. This filter absorbs yellow-green, green, blue-green, blue-violet and ultra-violet fairly heavily. Its transmission value for tungsten light is 34%.

Filter F_3 —Ilford no. 302 Minus Red. This filter absorbs orange and near red heavily although absorption decreases from 6800 Å. to infra-red. Transmits increasingly from 5900 Å. towards 4800 Å. with a secondary absorption maximum at 4100 Å. Transmission reasonably free from 4400 to 5400 Å. Transmission value for tungsten light 20%.

Filter F_2 —depending upon the experimental group this is the same as either F_1 or F_3 .
Test patch (as seen by left eye)—visual angle subtended $1^\circ 54'$, brightness 1.8 lfs. Brightness of background 1.9 lfs. Right eye background brightness 2.9 lfs. Angle subtended by stimulus words $2^\circ 24'$.

Subjects

Eight males, four females, all members of Department of Psychology at University College, London. All showed normal colour vision as indicated by the Ishihara test.

Procedure

After 5 min. adaptation the subject was asked to look into the stereoscope and report upon the colour of the test patch. The action of the lever was explained to him: he was told that if he pushed the lever forward the spot would gradually become red and if he pulled it back it would gradually become green. Between these two extremes he would see it as colourless.

He was then instructed to manipulate the lever so that he alternately saw the spot as 'just going red' and 'just going green'. Finally, he was asked to indicate (by pressing the button) during any traverse from red to green or vice versa that point at which the spot became grey or colourless.

After several practice trials the spot was set at the colourless point and the subject's right eye threshold found for a line of light which subtended the same angle as the stimulus words. Threshold determinations were made by the method of limits using descending presentations only. The awareness threshold was taken as that value not equalled or improved upon for the three following presentations. For these threshold determinations half the subjects, group A, saw the bar of light in red, the other half, group B, in green. For both groups the light intensity for all subsequent presentations was set at 0.2 log units below the determined threshold.

The subject was then instructed to proceed as before with controlling the spot. Each subject was given four trials of five traverses from red to green and back again within each trial. During the first trial, which was a practice, no stimulus presentation was made to the right eye. In one of the subsequent trials, the subject was presented subliminally with the word CANCER; in another with the inverted mirror image of CANCER. The third trial was again with no stimulus presentation. The order of presentation of these trials was different for each individual. Thirty seconds rest was allowed between trials. At the conclusion of the last trial each subject was asked whether he had seen anything through the stereoscope apart from the colour test patch. None of the subjects reported seeing anything.

RESULTS

The curves in Figs. 2 and 3 show the changes in threshold for red and green and the change in the relative amount of red and green to produce grey, as a function of time, for the two groups.

It will be seen that, in respect of thresholds recorded for the 'emotional' and 'neutral' words, respectively, the curves of the one group are very nearly an inversion of those for the other.

For the purposes of a statistical test of this difference the thresholds on the first two responses in each trial have been compared with those on the last two responses. In the case of the 'red' group five out of the six subjects showed an initially higher threshold for red and lower threshold for green when E words as opposed to N words were presented, followed by the reverse effect wherein the threshold for red fell while that for green rose. Using a two-tail test the differential change in the red threshold for the two sorts of word for the A group is significant at the 4 % level of probability ($t = 3.95$). The change in the threshold for green for the same group was also significant ($t = 3.18$).

In the case of B group for whom subliminal stimulation was in 'green' the differential change in the green threshold fell just short of that required for statistical significance ($t = 2.16$). For this group the parallel change in the red threshold also fell short of that required for statistical significance.

For both groups the difference between the thresholds for the two sorts of stimulus material was greater for the colour with which they were subliminally stimulated.

The difference between the two groups in respect of the direction of change in threshold for the two colours is significant at the 2 % level of confidence in the case of both the 'red' and 'green' thresholds ($t = 3.18$ and 3.78 , respectively, with two-tail tests). In both groups five out of six subjects showed the changes remarked upon above. By the sign test (two-tail) this is significant ($P < 0.04$).

Along with the changes in threshold for the two colours there went a change in the ratio of red to green for the perception of grey. Whereas the 'red' group showed a significant ($t = 3.498$, $P < 0.02$) increase in the amount of red for 'N' words and decrease in the case of 'E' words, the 'green' group showed the reverse effect—the difference between the groups being significant at the 2 % (two-tail) level of confidence ($t = 3.58$).

Summarizing the findings so far it can be said that whereas the red group showed during presentation of the 'emotional word' a generally lower but progressively increasing sensitivity to red with decreasing sensitivity to green, the other group showed a generally higher and increasing sensitivity to green with an insignificant decrease in sensitivity to red.

This apparent anomaly, wherein the one group showed a lower sensitivity to the colour component with which they were subliminally stimulated, whereas the other group showed a raised sensitivity to the colour with which they were stimulated, is consistent with the results from a second experiment which will be referred to again, later in this paper.

If the changes in threshold resulted from a change in sensitivity of the peripheral receptor system, then it might be expected on *a priori* grounds that the changes in

the ratio of the two component colours to produce the subjective experience of grey would parallel the changes in threshold for the individual components. A glance at Figs. 2 and 3 shows this to be in fact the case when the subliminal stimulation was green but not when it was red.

In other words the green group showed not only a heightened sensitivity to green when the E word was shown but also under these conditions relatively less green in the red/green achromatic mixture. The red group, on the other hand, showed a very much raised threshold for red, but less red than green in the achromatic mixture during stimulation of the other eye with the E word.

This difference between the groups points to another fact with certain theoretical implications.

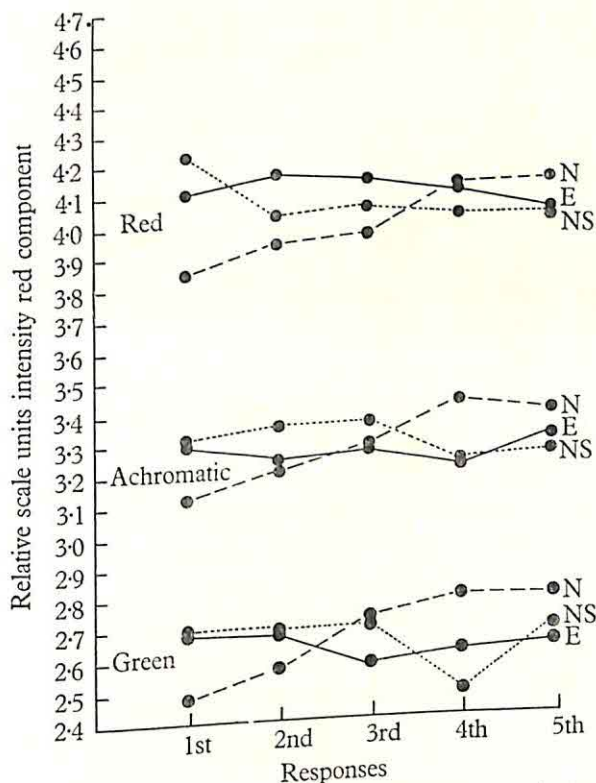


Fig. 2. Experiment I, group A, stimulus to right eye red, six subjects. E = Emotional stimulus, N = neutral stimulus, N.S. = no stimulus.

The centre curves on Figs. 2 and 3 which show the achromatic mixtures for the two sorts of stimulus material under the respective colour conditions are in fact means derived from the raw data curves shown in Figs. 4 and 5. Examination of these latter shows that in the case of the 'red' group the point signalled by the subject as achromatic when traversing from red to green was much further from the red threshold during stimulation with an E word than with an N word.

With the 'green' group, on the other hand, the tendency was reversed. In this case when traversing from red to green the point of achromaticity was signalled earlier during subliminal stimulation with the E word than with the N word.

When traversing from green to red, however, the subjects in the (green) group B

went further in the red before signalling achromaticity during presentation of the E word than when the N word was presented.

In other words, having just seen the test spot as the same colour as that with which they were in fact subliminally stimulated, both groups showed a tendency to go on seeing it coloured longer when the stimulus to the other eye was an E word than when it was an N word. Five out of six subjects in group A and four out of six in group B contributed to this effect. For both groups together this finding is significant at the 2% level of confidence by a two-tail test ($t = 2.67$).

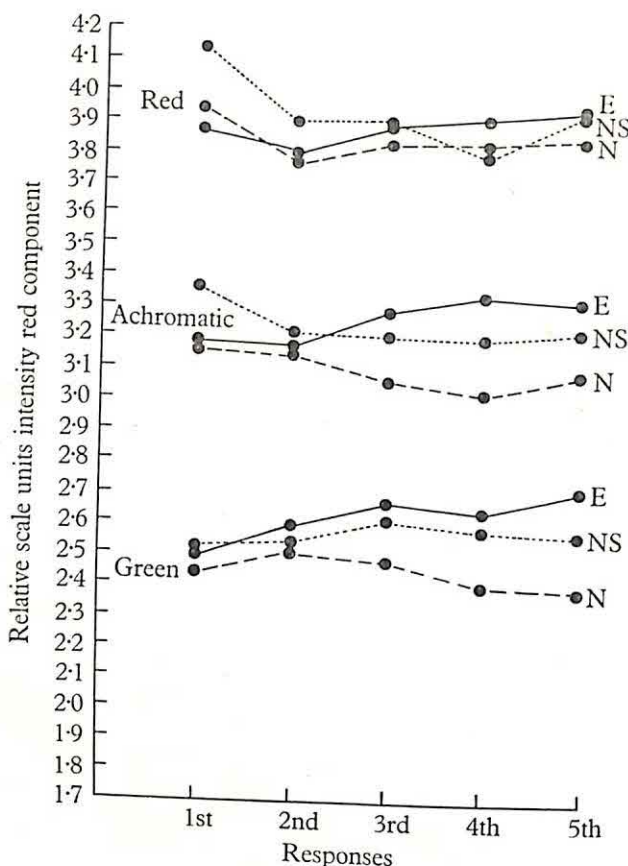


Fig. 3. Experiment I, group B, stimulus to right eye green, six subjects. E = Emotional stimulus, N = neutral stimulus, N.S. = no stimulus.

The relative effect of applying no subliminal stimulus to the right eye may be seen in Figs. 2 and 3. The effect appears, generally, to be somewhere between that evoked by the two sorts of subliminal stimulation.

One implication of this finding is that, irrespective of the meaning it connotes, a subliminal stimulus to one eye has an effect upon the sensitivity of the other eye.

The most striking occurrence of this tendency was in the (green) B group. All the subjects in this group showed a greater difference between the green thresholds for E and N words, respectively, than they did between the thresholds recorded for either of these words and those recorded when 'no stimulus' was exposed to the other eye.

By the sign test this result is significant at the 5% level of confidence (two-tail).

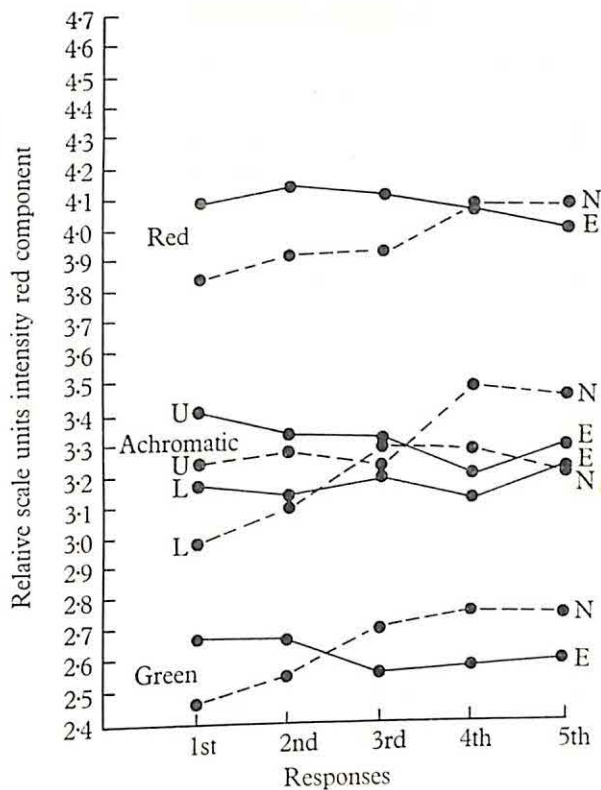


Fig. 4. Experiment I, group A, stimulus to right eye red, six subjects.
E = Emotional stimulus, N = neutral stimulus.

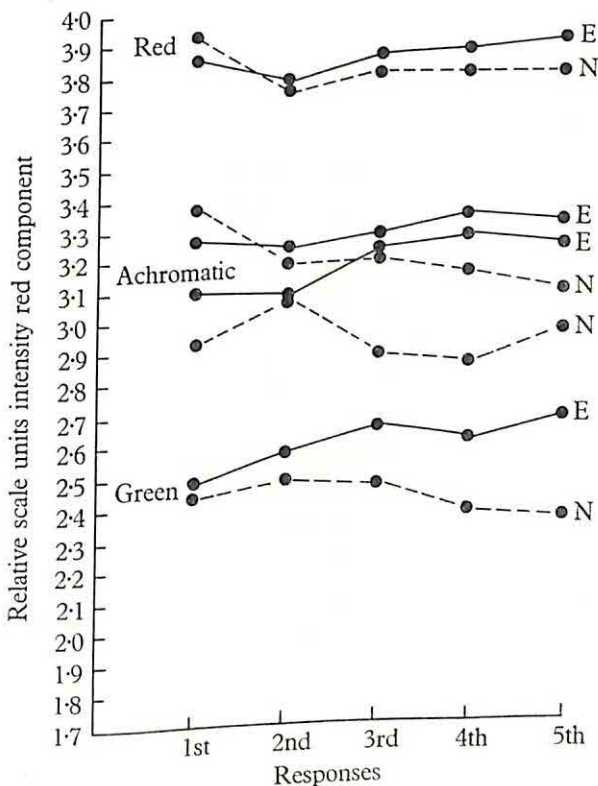


Fig. 5. Experiment I, group B, stimulus to right eye green, six subjects.
E = Emotional stimulus, N = neutral stimulus.

DISCUSSION

This study was based upon an earlier research which suggested that the visual brightness threshold of one eye could be changed by subliminal presentation of different sorts of stimulus material to the other eye.

The method was designed to discriminate between central and peripheral processes as likely factors in threshold change. In particular it investigated colour threshold changes, as a function of time of response, for two sorts of stimulus material, the effect upon these changes of using different wavelengths for the subliminal stimulus, and the relationship between colour threshold changes and changes in the relative amounts of these colours for the subjective experience of grey in the test patch.

The results are consistent with those from the earlier experiment to the extent that they suggest that the emotional connotations of a stimulus presentation to one eye will affect the threshold recorded for the other eye. They suggest, moreover, that those wavelengths for which the threshold changes are determined by the wavelength of the subliminal presentation.

The slopes of the different threshold curves suggest that whatever their initial threshold subjects become increasingly sensitive to perception of the colour in which the 'emotional' stimulus item is being presented and relatively less sensitive to that component in which the neutral stimulus is presented.

An anomalous finding is that, in terms of the colour of the subliminal stimulus, threshold changes differed for the two groups. Whereas the red group were generally less sensitive to red during the presentation of E words the green group were significantly *more* sensitive to green during the presentation of the 'emotional' subliminal stimulus item.

A possible explanation of this finding is that whether the response to anxiety material will be one of vigilance or one of defence might well be a function of other cues presented simultaneously. In this case the 'other cues' could be the two colours red and green. That these results are not fortuitous is supported by the findings from a second subsidiary experiment to which reference was made earlier. This experiment differed from that already described in only one respect. Instead of the neutral stimulus consisting of the negative image of the anxiety word it consisted of the neutral word 'stance', the word which in Haider's experiment, described earlier, evoked the smallest rise in threshold.*

For the experiment referred to half the subjects had the words CANCER and STANCE in green, the other half in red. The threshold curves are shown in Figs. 6 & 7 respectively.

While there are considerable differences between these curves and those for the other experiment, it will be seen that the main results were the same, namely, that during the presentation of the anxiety word the red group showed a raised threshold for red, while the green group showed a very much lowered threshold for green. The rise in the red threshold for the red group and fall in green threshold for the green group were greater than the rise in red threshold for the 'green' group and fall in green threshold for the red group.

* These results are consistent with Kravkov's (1941) finding that adrenaline, and such indirect stimulation as acts sympathico-mimetically, decrease visual sensitivity to red-orange light and increase it to green-blue light.

The question remains as to whether the observed threshold changes reflect a variation in peripheral sensitivity or some more central change.

While the differences between changes in threshold for the component colours and changes in the ratio of these components for experiencing the test patch as achromatic are not sufficient to exclude the possibility of variation in peripheral sensitivity, there is some evidence for a more centrally placed determinant of threshold change.

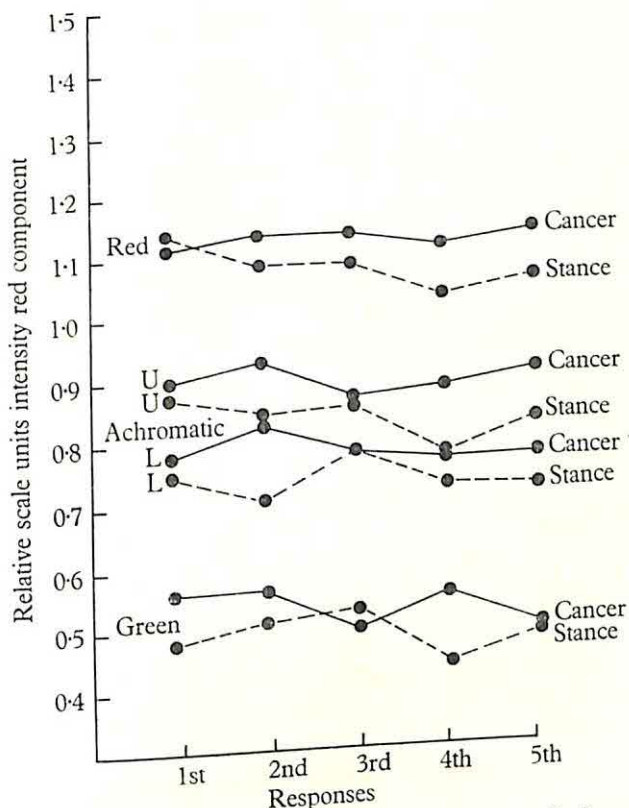


Fig. 6. Experiment II, group A, stimulus to right eye red, six subjects.

In the first place the observed changes were greater for the colour with which the subject was subliminally stimulated. If the observed changes were a function of alterations in peripheral sensitivity, then any change in threshold for one of the colours should have been accompanied by an equal and opposite change in the complementary colour threshold.

In the second place the variations in the achromatic range are extremely difficult to explain as a function of peripheral change but understandable in terms of decision processes.

Apropos of this the evidence suggests the possibility that the subliminal stimulus operates upon the probability matrix underlying cognitive processes, thus delaying or advancing recognition of those classes of events which have characteristics in common with the 'emotional' stimulus.

Such an eventuality does not of course imply voluntary suppression in the sense used by critics of the traditional perceptual defence experiment. Since the subject is never aware of the stimulus item, nor required to make a verbal report, there is no

reason to suppose that he would voluntarily delay making the quite innocuous response of moving the control lever.

In conclusion let it be said that this was a speculative exercise in the development of a technique to distinguish psychologically between the operation of certain hypothetical neural mechanisms. Manifestly, it has raised more questions than it has answered. If nothing else it has at least demonstrated the need for further research in this area.

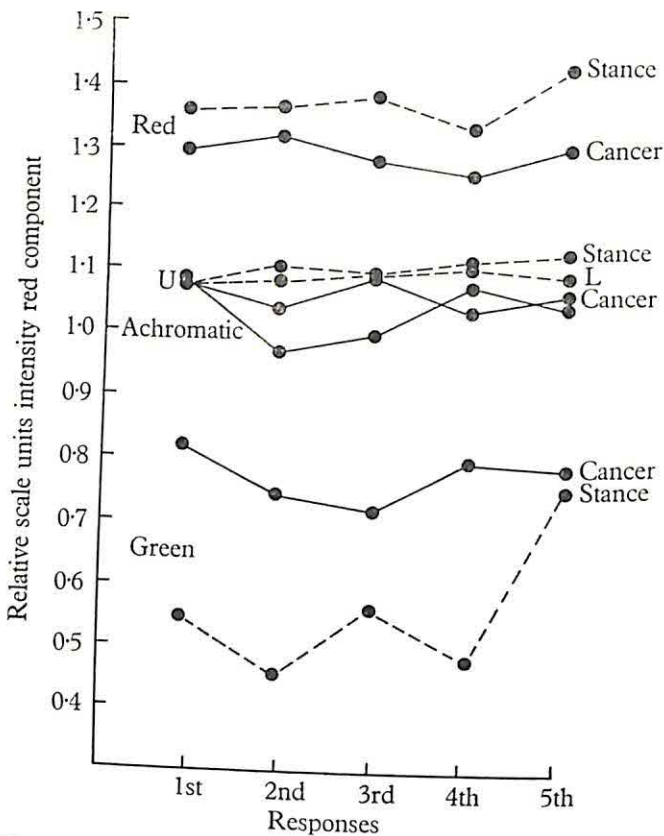


Fig. 7. Experiment II, group B, stimulus to right eye green, five subjects.

REFERENCES

- BABSKII, E. B. (1943). The significance of the sympathetic nervous system in regulation of excitability of the visual analyser. *Probl. Fiziol. Opt.* **4**, 17-30.
- BENDER, M. B. (1952). *Disorders in Perception*. Springfield, Ill.: Thomas.
- BERNHOUT, M., GELLHORN, E. & RASMUSSEN, A. T. (1953). Experimental contributions to the problem of consciousness. *J. Neurophysiol.* **16**, 21-36.
- BLUM, G. S. (1955). Perceptual defense revisited. *J. Abnorm. (soc.) Psychol.* **51**, 24-9.
- BROADBENT, D. E. (1958). *Perception and Communication*. London: Pergamon.
- DIXON, N. F. (1958a). Apparent changes in the visual threshold as a function of subliminal stimulation. *Quart. J. Exp. Psychol.* **10**, 211-9.
- DIXON, N. F. (1958b). Apparatus for continuous recording of the visual threshold by the method of 'closed loop control'. *Quart. J. Exp. Psychol.* **10**, 62-3.
- DODT, E. (1956). Centrifugal impulses in rabbit's retina. *J. Neurophysiol.* **19**, 301-7.
- FRENCH, J. D. & HERNÁNDEZ-PEÓN, R. (1955). Projections from cortex to cephalic brain stem in monkeys. *J. Neurophysiol.* **18**, 74-95.

- FRENCH, J. D., LIVINGSTON, R. B. & HERNÁNDEZ-PEÓN, R. (1953). Cortical influences upon the arousal mechanism. *Trans. Amer. Neurol. Ass.* **78**, 57-9.
- GANGLOFF, H. & MONNIER, M. (1956). Electrographic aspects of an arousal or attention reaction induced in the unanaesthetized rabbit by the presence of a human being. *Electroenceph. Clin. Neurophysiol.* **8**, 623-30.
- GELLHORN, E. (1954). Physiological processes related to consciousness and perception. *Brain*, **77**, 401-15.
- GRANGER, G. W. (1957). Night vision and psychiatric disorders. A review of experimental studies. *J. Ment. Sci.* **103**, 48-79.
- GRANIT, R. (1955). Centrifugal and antidromic effects on ganglion cells of the retina. *J. Neurophysiol.* **18**, 388-411.
- HOWIE, D. (1952). Perceptual defence. *Psychol. Rev.* **59**, 308-15.
- KRAVCOV, S. V. (1941). Colour vision and the autonomic nervous system. *J. Opt. Soc. Amer.*, **31**, 335-7.
- INGVAR, D. H. & HUNTER, J. (1955). Influence of visual cortex on light impulses in the brain stem of the unanaesthetized cat. *Acta physiol. Scand.* **33**, 194-218.
- LINDSLEY, D. B. (1956). Physiological psychology. *Annu. Rev. Psychol.* **7**, 323-48.
- PETTIGREW, T. F., ALLPORT, G. W. & BARNETT, E. O. (1958). Binocular resolution and perception of race in South Africa. *Brit. J. Psychol.* **49**, 265-78.
- ROTHBALLER, A. B. (1956). Studies of the adrenaline sensitive component of the reticular activating system. *Electroenceph. Clin. Neurophysiol.* **8**, 603-22.
- SAMUELS, I. (1959). Reticular mechanisms and behaviour. *Psychol. Bull.* **56**, 1-25.
- THERMAN, P. O. (1938). The neurophysiology of the retina in the light of chemical methods of modifying its excitability. *Acta Soc. Sci. fenn., NSB*, II, 1.

(Manuscript received 27 November 1959)

THE EFFECTS OF STIMULANT AND DEPRESSANT DRUGS ON THE BIDWELL PHENOMENON

BY SATORU AIBA

Institute of Psychiatry, University of London

The effects of amylobarbitone and dexamphetamine on the Bidwell phenomenon were studied. The measure used as the dependent variable was the intensity of the red coloured stimulus which was necessary if it were to become just visible when followed by a fixed white stimulus. It was found that the intensity threshold of the red stimulus was raised under amylobarbitone and lowered under dexamphetamine. There was a difference in the latent period of the action of these drugs, dexamphetamine acting much more quickly than amylobarbitone. Possible physiological mechanisms responsible for the production of the phenomenon are discussed and the results of the present experiment evaluated in the light of available evidence concerning the action of drugs on these processes.

I. INTRODUCTION

If a coloured light stimulus is viewed through a sector disk such as the one illustrated in Fig. 1, rotating in the direction of the arrow, the stimulus becomes imperceptible, and instead its negative after-image is seen in the place of the original stimulus. In order to obtain an adequate effect, the disk must be sufficiently illuminated and the speed of rotation be somewhere in the region of 5 r.p.s. When it is seen under favourable conditions, it gives the observer the impression that the negative after-image is the stimulus being presented, for the original stimulus cannot be seen at all. The phenomenon was first demonstrated by Bidwell (1896), and since then it has been frequently referred to in text-books of psychology and ophthalmology.

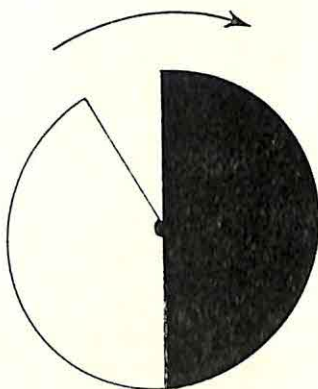


Fig. 1. Bidwell's disk.

Eysenck & Aiba (1957), in a series of experiments instigated by Lehmann's preliminary report in 1950, studied the effects of drugs on this phenomenon. They assumed that the phenomenon is the result of the eye's response to the coloured stimulus being inhibited by the white light that followed it. In order to get a measure of this inhibitory effect of the white light, they measured the highest intensity of the

coloured stimulus that a given white light could suppress completely. (We shall refer to this critical intensity of the coloured stimulus as the *threshold* of the coloured stimulus.) In one of their experiments, Eysenck & Aiba found that the threshold of the coloured stimulus was raised by amylobarbitone and lowered by dexamphetamine. In another experiment, however, they found that both these lowered the threshold. The main difference between the conditions of the two experiments was that, whereas in the first case the duration of the succeeding white stimulus was relatively long (125 msec.), in the second experiment it was generally much shorter (40–120 msec.). The authors inferred from this that the discrepancy of the effects of the drugs was due to this difference in conditions. They concluded that the main effect of amylobarbitone is to raise the threshold of the coloured stimulus and that of dexamphetamine to lower it, provided the duration of the succeeding white stimulus is equal to or longer than 125 msec. This conclusion is still tentative in view of the presentation of an additional white stimulus *before* the coloured stimulus which occurred in the first experiment, but not in the second one. Although the change in the length of intervals between this additional white light and the coloured stimulus did not have any apparent effect (see Eysenck & Aiba, 1957, fig. 1), the results of the first experiment might well have been different without this additional white stimulus.

The present experiment was undertaken to test further Eysenck & Aiba's hypothesis regarding the main effects of amylobarbitone and dexamphetamine, in what were thought to be optimal conditions for that purpose. Thus, the design had the following features: (1) a longer succeeding white stimulus was used to ensure that the unknown factor accompanying shorter durations was eliminated; (2) the additional preceding white light was eliminated; (3) pupillary control was introduced, since it was thought that the drugs might affect the size of the pupil and bring in some undesirable errors in the measurement.

II. THE EXPERIMENT

Apparatus. The apparatus used for this experiment is virtually identical with that used in the experiments of Eysenck & Aiba (1957). A fuller description of it appears in this article and a photograph in Eysenck, 1957. For this particular study the apparatus was set to give the following sequence of events: (i) a dark interval lasting approximately 3.5 sec., (ii) a coloured stimulus (red), surrounded by a white annulus, lasting 10 msec., (iii) a white stimulus lasting 350 msec. immediately following the red stimulus. The luminance of the red stimulus, which was the dependent variable in the experiment, could be varied between 0 and 2000 ml. (approx.). The spectrum composition of it was approximately 650 m μ into infra-red. The luminances of the white stimulus and the white annulus surrounding the red stimulus were fixed both at 95 ml. (During the dark interval, the field became about a hundredth of the luminance of the white stimulus, i.e. approximately 1 ml.) Both the coloured stimulus and the white stimulus were centrally fixated. The red circular area in the coloured stimulus subtended a visual angle of 1°, while the outer circumference of the white annulus and the succeeding white stimulus subtended a visual angle of 4°. The stimuli were viewed through a 3 mm. artificial pupil.

Subjects. Five subjects were used in the present experiment, all of whom were students in a post-graduate course of clinical psychology. The ages of the subjects ranged between 21 and 33, with a mean of 24. None had any gross defect in vision and all had a visual acuity of 6/9 or better as tested by a Snellen chart. There were three men and two women.

Procedure. After 15 min. of dark adaptation, the subjects were told to look at the centre of the stimulus patch indicated by means of a fixation spot and to increase the intensity of the red

stimulus by turning the knob of a variac, until they could see the first suggestion of red in it. (In the Eysenck & Aiba study, thresholds were determined by descending series—the subject brought down the intensity of the red stimulus till he no longer saw red, but we later found that ascending series generally gave more reliable results.) The subjects were allowed to repeat this if they were in doubt on the first occasion. As soon as they were satisfied with their adjustment, they resumed their dark adaptation by staying in the dark until the next trial began. Four of these trials made up one session, and during the next 2 hours the same procedure was repeated every 20 min. Thus, there were altogether six sessions, each consisting of four trials, and between each session there was an interval of 5–7 min. depending on how quickly the subjects made each adjustment. During these intervals also they stayed in the dark. Owing to the great difficulty experienced by the subjects in keeping the eye close to the eye-piece and in maintaining correct fixation during the lengthy course of the experiment, a certain amount of prior training was necessary. This consisted of repeating part of the experimental procedures, i.e. the first, the second and the third sessions, on three different days prior to the experiment proper; consequently, the subjects were thoroughly familiarized with the experimental procedure when they entered the experiment.

Drugs. The drugs were administered between the second and the third sessions each day. The drugs and their doses were identical with those used in the Eysenck and Aiba experiments, i.e. (1) dexamphetamine sulphate 10 mg. (Dexedrine, Smith Kline and French); (2) amylobarbitone sodium 290 mg. (Sodium Amytal, Lilly), and (3) a placebo. These were given in identically appearing capsule form and taken orally. Each subject was given the drugs in random order over the 3 days of testing. Between experimental days at least a week elapsed, so the sequence effect due to the residual drug effects was negligible.

Results. A preliminary analysis of the data had disclosed that there is a large day-to-day variation in the basic levels of threshold readings for each subject, and also that there is no systematic trend in that variation. Consequently, it was decided to base the analysis only on relative values of the readings within each day. For this purpose the average of the four readings of the first session was chosen as a reference point for that particular day, and this was subtracted from readings of all subsequent sessions of that day. Moreover, only the average of the four trials in each session was eventually used for the analysis, for, as the four trials are correlated, including them as independent degrees of freedom would tend to over-emphasize the results of the experiment.

The outcome of the analysis of variance is shown in Table 1. As can be seen from this, the *F* tests based on all sources of variation are at a highly significant level. Of particular importance are the variations due to drugs and those due to the interaction between drugs and sessions. These show that the drugs

Table 1. *Analysis of variance for the results of the different sessions (sessions II–VI)*

Source	D.F.	S.S.	M.S.	<i>F</i>
Between people	4	12,206	3,051.5	25.39***
Between sessions	4	5,482	1,370.5	11.40***
Between drugs	2	2,451	1,225.5	10.20***
Interaction				
Between sessions/people	16	2,834	177.1	5.35***
Between drugs/session	8	2,022	252.8	7.63***
Between people/drugs	8	1,776	222.0	6.71***
Residuals	32	1,059	33.1	—
Total	74	27,830	—	—

*** significant at 0.1 % level.

produced definite effects, but that the time at which the effects of the drugs became manifest was not the same. How each drug altered the threshold in course of time can be seen from Table 2. Differences between average readings under different treatments were tested for their significance for each session and the results shown in Table 3. It may be seen from them that there are no significant differences between the treatments either in session II or in session III. In sessions IV and V, however, there is a significant difference between the placebo and dexamphetamine and also between amylobarbitone and dexamphetamine (both at 1%), but there is no significant difference between the placebo and amylobarbitone in either of these sessions. In session VI, our last session, the difference between amylobarbitone and the placebo becomes significant (at 1%) and the differences between dexamphetamine and the placebo, and between dexamphetamine and amylobarbitone still remain significant (at 5 and 1% respectively). As to the direction of the effects of the two drugs, they were in the opposite direction, i.e. dexamphetamine lowered the threshold while amylobarbitone raised it.

Table 2. *The average reading for each drug in each session*

Session Drugs]	II	III	IV	V	VI
Placebo	3.05	5.50	8.10	8.90	7.85
Dexedrine	2.00	4.40	3.30	5.10	5.60
Amytal	3.35	3.80	8.30	8.65	12.95

Table 3. *T-tests for the differences between average readings for the drugs in each session*

	Direction	t	Sig. level
Session II	Pl > Dx	1.154	N.S.
	Pl < Am	0.330	N.S.
	Am > Dx	1.484	N.S.
Session III	Pl > Dx	1.209	N.S.
	Pl > Am	1.869	N.S.
	Am < Dx	0.660	N.S.
Session IV	Pl > Dx	5.277	**
	Pl < Am	0.220	N.S.
	Am > Dx	5.496	**
Session V	Pl > Dx	4.177	**
	Pl > Am	0.275	N.S.
	Am > Dx	3.793	**
Session VI	Pl > Dx	2.473	*
	Pl < Am	5.606	**
	Am > Dx	8.080	**

* Significant at 5% level. ** Significant at 1% level. N.S. = not significant.

III. DISCUSSION

Our findings on the effect of dexamphetamine and amylobarbitone on the threshold of the coloured stimulus are in agreement with the result of Eysenck & Aiba's first experiment. Thus, as far as the effects of these drugs are concerned, the simple sequence of stimuli 'dark-red-white' is not different from the more complex one

'white-dark-red-white' provided the white stimulus following the red is longer than 125 msec.

In this study we obtained new information about the latent period of the effects of the drugs on Bidwell's phenomenon. Considering the above reported results in terms of the time from the administration of the drugs, we find that the threshold was lowered by dexamphetamine within 30 min. from the administration of the drug and it remained so for the rest of the testing time (approximately 40 min. from when it first became apparent); on the other hand, with amylobarbitone, the effect did not appear till about 70 min. after the administration. These figures agree fairly well with those obtained in a study dealing with the effects of drugs on critical frequency of flicker (C.F.F.). Roback, Krasno & Ivy (1952) found that the peak effect of dexamphetamine (10 mg.) on C.F.F. occurred at about 30 min. from the time of administration, and that of quinalbarbitone (96 mg.) between 30 and 75 min.

Before considering the mechanism of the effects of drugs on the threshold of the coloured stimulus, it might be worth while noting that in Eysenck & Aiba's second experiment the duration of the succeeding white stimulus was mostly below 100 msec., whereas in the present experiment (and in Eysenck & Aiba's first experiment) it was much longer. Now, in psychophysics, 100 msec. is usually regarded as the upper limit of the duration where reciprocal relationships between time and intensity exist (the Bunsen-Roscoe law). It is reasonable to suppose that in our phenomenon, too, these relationships exist for the succeeding white stimulus, provided it does not exceed 100 msec. Thus, an increase in the *duration* of the white stimulus should have the same effect as an increase in the *intensity* of that stimulus, i.e. a higher threshold for the coloured stimulus (the opposite would happen if there is a decrease in the duration). This, of course, is the result of temporal summation of the effect of the white stimulus.

It is possible that a drug may affect this temporal summation. If the drug decreases the efficiency of temporal summation, a white succeeding stimulus, of a given duration and intensity, would behave as if it were of shorter duration and hence of a lower intensity. The opposite would happen if the drug increases it. However, these phenomena will occur *only* when the duration of the white stimulus is within the limits of temporal summation. It would not be surprising, therefore, if we had found different results in an experiment involving a light stimulus much longer than 100 msec., from what we obtained in an experiment involving a light stimulus shorter than 100 msec.

How then would the effects of drugs be accounted for, when the duration of the succeeding white stimulus is far above 100 msec. as in the present experiment? In this condition, the change in duration of the white stimulus would have no influence on the threshold of the coloured stimulus, and the *sole* determining factor would be its intensity. One may be inclined to infer from this that the effects of drugs on the threshold are through their effects on the eye's response to the white stimulus, causing a change similar to that which variations in its *intensity* might have brought about. Here, we are in need of two kinds of evidence if the above inference is to be justified. One would be evidence for any physiological process in the visual system which might be relevant to our phenomenon and is directly influenced by the intensity of a stimulating light, and the other evidence of the existence of any drug that would

cause exactly the same effect in that process *without* the actual change of the *physical intensity* of the stimulus. We believe that we are in possession of both these types of evidence, though it does not necessarily follow from this that the effects of drugs are explainable only on the basis of the facts to be presented. For instance, more central types of explanation are possible, though not discussed for lack of concrete evidence.

One of the responses of the retina to a light stimulus is to set up an inhibitory potential before the excitation proper starts. The potential is inhibitory in as much as it is capable of inhibiting excitation, whether the excitation be the result of some previous stimulation or of a spontaneous activity of the retina. This is *pre-excitatory inhibition*, demonstrated by Granit (1947). Attempts have been made to correlate some *sensory* phenomena in human vision, such as alpha-adaptation and meta-contrast, with this process (e.g. Wright & Granit, 1938). There is a reasonably good correspondence between the size of *a*-waves in the electroretinogram (e.r.g.) and the extent to which this inhibition is active. So, one could use the former as an index of the latter. According to this index, pre-excitatory inhibition is a function of the intensity of a stimulus light, i.e. it increases with an increase in the intensity of the stimulus.

Certain drugs seem to alter the response of the retina to a light stimulus in the same way as an increase of stimulus intensity. Danis (1956) reports that the e.r.g. of rats is increased in amplitude in all component responses by pento- and phenobarbitone. This closely resembles what happens when the intensity of a stimulus is increased. Of particular interest to us is the increase in the size of the *a*-wave, for this indicates, according to the above criterion, the increase of pre-excitatory inhibition under these drugs. It is not surprising if amylobarbitone produces the same effect in these animals, for it belongs to the same group of drugs as the others (i.e. barbiturates). Nor is it entirely unlikely that the human retina behaves in the same way under these drugs as that of the rat, for they are both of the same type ('E-retina') (Granit, 1947).

Thus, the effect of amylobarbitone on the threshold of the coloured stimulus might become comprehensible if it could be shown that the phenomenon is dependent on pre-excitatory inhibition. There is, of course, no direct proof for this, but the available facts do suggest that it is the case.

First, suppression of the coloured stimulus never takes place, if the stimulus is presented *after* or *during* the white light; it occurs only when the coloured stimulus is *immediately followed* by the white flash. Nor does it occur when another white light is presented just before the coloured stimulus in such a way that the two white lights would summate. (If an additional white stimulus is given 70 msec. or so before the coloured stimulus, the phenomenon is observed even to a slightly increased extent—e.g. compare figs. 1 and 2 of Eysenck & Aiba. This is to be distinguished from an additional light being presented immediately before the coloured stimulus.) This all becomes understandable if we assume that pre-excitatory inhibition is responsible for the suppression of the coloured stimulus in our phenomenon. Boynton & Triedman's (1953) electrophysiological study shows that the *b*-wave is actually suppressed when the test flash causing it is followed by another flash. This could be taken as a manifestation of the effect of pre-excitatory inhibition, and as the conditions of this study resemble those under which the Bidwell phenomenon is observable, one may

infer that pre-excitatory inhibition is likewise responsible for the production of the phenomenon.

Further evidence that the phenomenon may be based on pre-excitatory inhibition is as follows. Light adaptation raises the threshold of the coloured stimulus in our phenomenon (Aiba, 1959). If the phenomenon were based on pre-excitatory inhibition, light adaptation should raise the threshold of the coloured stimulus as it did in Aiba's study, because pre-excitatory inhibition is a function of levels of light adaptation (Granit, 1947). (The gradual rise in the threshold of the coloured stimulus under all treatments, observed in the present experiment, might be due to a slow recovery from the initial dark adaptation, i.e. the testing contributing to light-adapt the eye.)

Thus, the phenomenon is probably dependent on pre-excitatory inhibition; if so, the rise of the threshold of the coloured stimulus with an increase of the intensity of the white stimulus is possibly due to increased pre-excitatory inhibition. There is, as we have said, some ground to believe that the drug enhances pre-excitatory inhibition. One may perhaps be justified in inferring from the above that the observed rise in the threshold of the coloured stimulus under amylobarbitone is due to enhanced pre-excitatory inhibition.

As regards dexamphetamine, there is as yet no study of the effect of this drug on the e.r.g. Therefore, there is no ground on which we can re-state our previous argument. If the same mechanism through which the effect of amylobarbitone is mediated operates with dexamphetamine also, one would expect that pre-excitatory inhibition is reduced by this drug. However, only alcohol (Granit, 1947) and hyperventilation (Alpern, Faris, Eskildsen & Garnett, 1955) have so far been found to have this effect. Similarities between these conditions and the condition produced by dexamphetamine are yet to be determined.

We have not yet considered the spatial factor involved in the Bidwell phenomenon. Nevertheless, that this factor is equally important can be easily demonstrated by making the sizes of the coloured and succeeding white stimuli equal; the phenomenon never becomes so marked under this condition as it does with a larger succeeding white stimulus. It will be noticed further that, in Bidwell's original demonstration as well as in Lehmann's study, the coloured stimuli were always presented on a white background (equivalent to our white annulus) but never on a black one. Our personal observation is that, when the coloured stimulus is surrounded by a black area, the part of it bordering the black area never becomes complementary. The white surrounding area, therefore, appears to play an important role in complete suppression of the coloured stimulus. Perhaps in the condition under which the phenomenon appears, the luminance of the white surrounding area is higher than that of the coloured stimulus, and, therefore, apparent luminance of that stimulus is reduced by simultaneous brightness contrast. Reduction of apparent luminance of a stimulus in simultaneous brightness contrast is due probably to inhibition of the 'on' response (the primary response) to that stimulus (Asher, 1950; Barlow, FitzHugh & Kuffler, 1957). It is not surprising, therefore, that the white surrounding area is so important in the complete suppression of the coloured stimulus.

There is some evidence that a barbiturate enhances this simultaneous contrast effect. Trent (1947) reports an increase of contour acuity in subjects under pento-

barbitone anaesthesia. This result may be interpreted as due to heightened contrast between the white background which he used and his 'contour' objects. Unfortunately, there is not yet any similar psychophysical study to confirm this finding, nor do we have any evidence about the effect of dexamphetamine in this connexion. However, it is a possibility that the drugs may affect the Bidwell phenomenon through their effects on brightness contrast sensitivity.

The author is indebted to his colleagues at the Institute of Psychiatry, especially to Prof. H. J. Eysenck, Drs G. W. Granger and H. Holland for their critical reading of the manuscript, and to Dr A. E. Maxwell for his statistical advice. It must be pointed out, however, that opinions expressed are those of the author.

REFERENCES

- AIBA, S. (1959). Effect of light adaptation on the Bidwell phenomenon. Unpublished study.
- ALPERN, M., FARIS, J., ESKILDSEN, P. & GARNETT, P. (1955). Effect of hyperventilation on the human electroretinogram. *Science*, **121**, 101-2.
- ASHER, H. (1950). Contrast in eye and brain. *Brit. J. Psychol.* **40**, 187-94.
- BARLOW, H. B., FITZHUGH, R. & KUFFLER, S. W. (1957). Change of organization in the receptive fields of the cat's retina during dark adaptation. *J. Physiol.* **137**, 338-54.
- BIDWELL, S. (1896). On subjective colour phenomena attending sudden change of illumination. *Proc. Roy. Soc. Lond.* **60**, 368-77.
- BOYNTON, R. M. & TRIEDMAN, M. H. (1953). A psychophysical and electrophysiological study of light adaptation. *J. exp. Psychol.* **46**, 125-34.
- DANIS, P. (1956). Modifications de L'électrorétinogramme du rat produites par l'injection intra-artérielle proche de potassium, de vératrine et de narcotiques. *J. Physiol. Paris*, **48**, 479-83.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge & Kegan Paul.
- EYSENCK, H. J. & AIBA, S. (1957). The effects of stimulant and depressant drugs on the suppression of the primary visual stimulus. *J. Ment. Sci.* **103**, 661-5.
- GRANT, R. (1947). *Sensory Mechanisms of the Retina*. London: Oxford University Press.
- LEHMANN, H. E. (1950). Preliminary report on a device for the objective measurement of the negative after-image phenomenon. *Science*, **112**, 199-201.
- ROBACK, G. S., KRASNO, L. R. & IVY, A. C. (1952). Effect of analeptic drugs on the somnifacient effect of seconal and antihistaminics as measured by the flicker fusion threshold. *J. Appl. Physiol.* **4**, 566-74.
- TRENT, S. E. (1947). Narcosis and visual contour clarity. *J. Gen. Psychol.* **36**, 65-78.
- WRIGHT, W. D. & GRANT, R. (1938). On the correlation of some sensory and physiological phenomena of vision. *Brit. J. Ophthal.* (Suppl.), no. ix.

(Manuscript received 28 July 1959)

EXTRAVERSION, REMINISCENCE AND SATIATION EFFECTS

By R. LYNN

Exeter University

Eysenck's theory that extraverts accumulate reactive inhibition quickly and that it dissipates in them slowly and his application of this theory to after-effects and reminiscence is made the basis of six predictions: using the spiral after-effect, there should be (1) a negative correlation between extraversion and duration of the after-effect; (2) a tendency for extraverts to see progressively less of the after-effect with repeated massed trials; (3) a tendency for extraverts to recover more in their perception of the after-effect after a period of rest; (4) a negative correlation between the duration of the after-effect and a measure of reminiscence using the inverted alphabet printing task; (5) a positive correlation between extraversion and reminiscence; (6) a tendency for extraverts to show more work decrement with massed practice on the inverted alphabet printing task. Using forty male university students as subjects, predictions 1, 2, 4 and 5 are confirmed at a statistically significant level, and predictions 3 and 6 show non-significant results in the predicted direction.

I. INTRODUCTION

In view of the recent criticisms of Eysenck's (1957) theory linking introversion-extraversion with reactive inhibition, figural after-effects and satiation effects, this paper presents the results of an independent investigation of some predictions from the theory. Essentially, the Eysenck theory postulates that extraverts accumulate reactive inhibition quickly and dissipate it slowly. This theory mediates a large number of predictions, e.g. that extraverts should show greater work decrement and more reminiscence, for which there is confirmatory evidence (Eysenck, 1957). Eysenck has also postulated that figural after-effects and satiation effects are affected by reactive inhibition and hence related to introversion-extraversion and evidence in support of this postulate has been published for the kinaesthetic after-effect and the Archimedes spiral satiation effect (Eysenck, 1957).

The experiment reported here concerns the relation between extraversion, reminiscence and the spiral satiation effect. Eysenck's theory of the spiral satiation effect is as follows: as the subject fixates the spiral reactive inhibition is generated and this interferes with the after-effect according to its strength. Further, the processes underlying the reversal phenomenon must themselves generate satiation and therefore curtail the after-effect. Both processes would tend to make extraverts experience the after-effect less than introverts. Although little direct evidence is available, two additional hypotheses of Eysenck provide deductions about which there is evidence relevant to this theory. First, it is postulated that depressant drugs increase reactive inhibition and hence curtail the after-effect, while stimulant drugs decrease inhibition and hence increase the duration of the after-effect; this deduction has been partially confirmed by Eysenck, Holland & Trouton (1957). Secondly, it is postulated that brain injury increases the tendency to generate reactive inhibition and hence brain injured subjects should have smaller satiation effects. This deduction has been confirmed by Price & Deabler (1955), although some doubt has been thrown on this finding as a result of the findings of Spivack & Levine (1959) and what would appear to be the important question of the site of the injury remains uninvestigated. As far

as reminiscence is concerned, Eysenck accepts Hull's theory that reminiscence is due to the dissipation of reactive inhibition; when Eysenck's own postulate that extraverts generate reactive inhibition quickly and dissipate it slowly is added to Hull's theory, the prediction is that extraverts should show greater reminiscence.

This theory has not been without its critics, notably Hamilton (1959) and Rechtschaffen (1958), who maintain that there is insufficient evidence in its support. Rechtschaffen presents evidence showing some positive association between extraversion, figural after-effects and reminiscence using inverted alphabet printing, but the correlations were too low for statistical significance. However, since Rechtschaffen used an unusually short rest interval in his reminiscence test, it is arguable that his experiment cannot be regarded as an adequate test of the theory.

The present paper reports an investigation of six predictions from the Eysenck theory, namely: (1) under certain conditions of administration the inverted alphabet printing task should yield reminiscence scores which correlate positively with extraversion; (2) extraverts should show more work decrement in inverted alphabet printing under conditions of massed practice; (3) there should be a negative correlation between extraversion and duration of the spiral after-effect; (4) with repeated presentation, extraverts should show a greater fall off in the spiral after-effect as a result of their tendency to generate reactive inhibition more quickly; (5) after a period of rest extraverts should show a greater recovery in seeing the spiral after-effect as a result of the dissipation of reactive inhibition; (6) the duration of the spiral after-effect and reminiscence should be negatively correlated.

II. THE INVESTIGATION

Subjects. The sample consisted of forty male university students living in a university hall of residence; volunteers for the experiment were asked for and the first forty to volunteer were taken as the subjects. All subjects fell in the age range 18-23 with the exception of one who was slightly older. On the Maudsley Personality Inventory the mean score for extraversion was 23.5, S.D. 8.42, compared with the norm for the general population of 24.9, S.D. 9.71; and for neuroticism 27.6, S.D. 9.27, compared with the norm for the general population of 19.9, S.D. 11.02. It is evident therefore that the sample was representative of the normal population on the introversion-extraversion dimension but on neuroticism it scored somewhat more neurotic. However, it has been shown that English university students as a whole score more neurotic than the general population by something like $\frac{1}{2}$ S.D. (Lynn, 1959) and it may therefore be concluded that the present sample was representative of English university students though not representative of the general population on the neuroticism dimension.

Procedure. The spiral used in this experiment was a 4-throw spiral of 180° similar to the one illustrated by Eysenck (1957) except that it was 10 in. in diameter; it rotated at 78 rev./min., and subjects were tested in daylight conditions. Subjects were seated 6 ft. from the spiral and instructed to fixate the central point. The after-effect was explained to the subjects, and they were asked to report when the apparent movement ceased. Subjects were given a practice trial of 40 sec. They were then given ten trials of 30 sec. each in 'massed' conditions, i.e. the spiral was rotated as

soon as the subject had made his judgement of the cessation of the after-effect. Following this a 2 min. rest was allowed, after which subjects underwent one more trial. The reliability of this test, calculated by the split-half method, was 0.86.

Subjects were then asked to answer the Maudsley Personality Inventory which gives a measure of introversion-extraversion and neuroticism.

Finally, subjects were asked to do the inverted alphabet printing task. Subjects were instructed to print the alphabet upside down as fast as they could for fourteen trials with massed practice. They were then given a 2 min. rest, and asked to print the alphabet once more.

III. RESULTS

The variables about which predictions were made were scored in the following way: (1) reminiscence: a score was obtained by subtracting the speed on trial 15 of the inverted alphabet printing from the mean of the speeds on the trials 8-14; (2) work decrement: speed on trial 14 of alphabet printing; (3) duration of spiral after-effect: the reported duration on the first trial; (4) decline of the spiral after-effect: a score was obtained by subtracting the time of the after-effect on trial 10 from that on trial 1; (5) recovery of the after-effect with rest: duration of after-effect on the eleventh (post-rest) trial minus duration of the after-effect on trial 10. The product-moment correlations between these variables and extraversion are shown in Table 1. To afford comparison with other studies, the product-moment correlations of neuroticism with some of these measures are also shown in Table 1.

Table 1. *Product moment correlations of measures of inhibition with extraversion and neuroticism*

Measure	<i>E</i>	<i>N</i>	Measure 3
1. Reminiscence	+0.42*	+0.29	-0.34*
2. Work decrement	+0.21	—	—
3. Duration of spiral after-effect	-0.43*	-0.13	—
4. Decline of spiral after-effect	+0.42*	—	—
5. Recovery of after-effect after rest	+0.18	—	—
Speed of alphabet printing, trial 1	—	+0.29	—

* Significant at 0.05 probability level.

Two extreme groups of 12 extraverts and 11 introverts were obtained by considering only the subjects whose scores fell 7 or more points away from the mean. Fig. 1 shows the mean scores for the two groups over the 11 trials on the spiral after-effect and illustrates the greater decrement of the extraverts with repeated presentation. Fig. 2 shows the mean scores for the two groups on the inverted alphabet printing task and illustrates the tendency of extraverts to slow down and to show greater reminiscence. The means and s.d.'s of these two extreme groups on the measures calculated are shown in Table 2.

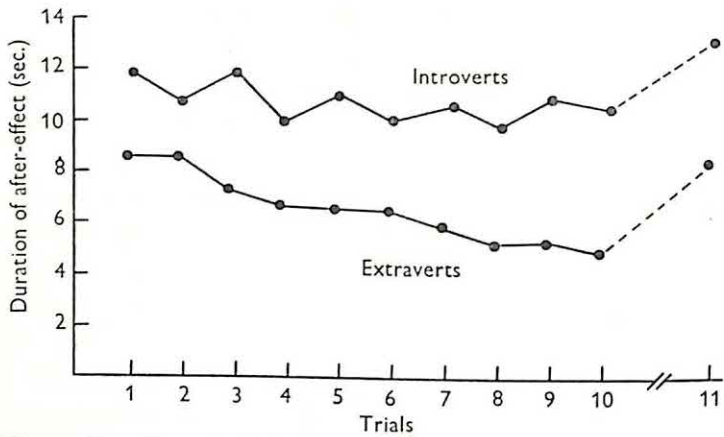


Fig. 1. Duration of spiral after-effects of introverts and extraverts

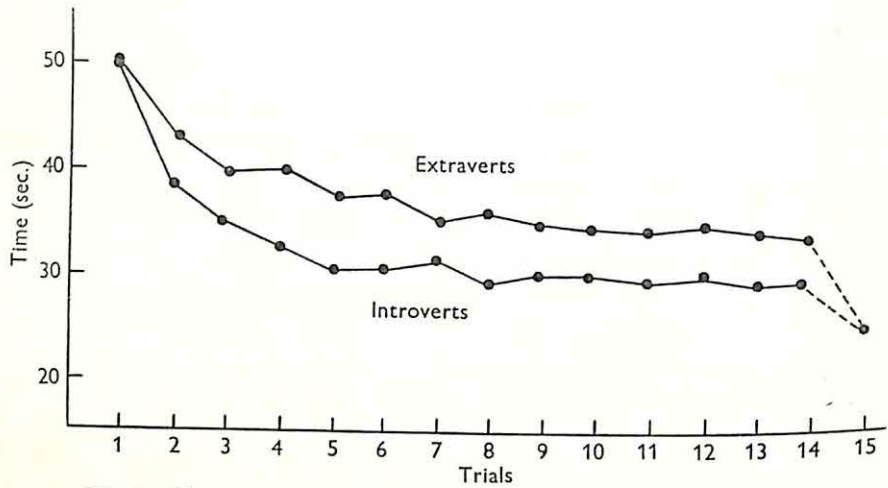


Fig. 2. Mean times and reminiscence of introverts and extraverts on the inverted alphabet printing task.

Table 2. Means and S.D.'s of introvert and extravert groups on measures of reactive inhibition

	Introverts		Extraverts	
	Mean	S.D.	Mean	S.D.
Inverted alphabet printing				
1. Reminiscence	5.10	2.28	10.42	4.98
2. Speed on trial 14	29.18	6.32	33.00	6.00
Spiral after-effect				
3. Duration on trial 1	12.09	8.50	7.52	5.31
4. Decline of after-effect	1.18	5.33	3.65	4.36
5. Recovery of after-effect	2.55	2.62	3.17	4.32

DISCUSSION

It is evident that the correlations tend to support Eysenck's theory. The two that do not reach statistical significance are in the predicted direction and the lowest, that of +0.18 between extraversion and recovery of the spiral, is probably reduced by the method of scoring, since the extraverts show a greater proportional post-rest

increase in after-effect which is to some extent masked by considering the absolute increase. On the other hand, it should be noted that the correlation between extraversion and the decline of the after-effect might have been increased by the experimental procedure used, since extraverts experienced more 'massing' as a result of their shorter after-effects.

The difference between the present results and those obtained by Rechtschaffen (1958) deserves some discussion. Rechtschaffen used Guilford's R scale for assessing extraversion, and this correlates only about 0.5 with the extraversion scale of the M.P.I. (Eysenck, 1959). But it is likely that the differences in methods of assessing reactive inhibition are also important. The method employed by Rechtschaffen was to give subjects five trials of distributed practice followed by five trials of massed practice and to obtain an estimate of reactive inhibition by subtracting the number of letters printed under massed practice from the number printed under distributed practice. Using this measure he obtained a correlation of +0.18 between extraversion and the amount of reactive inhibition accumulated. He also obtained a score of reminiscence by giving a one minute rest after the five massed trials, followed by two further trials. Reminiscence and extraversion correlated only +0.08. The difference between this finding and the present one might be due to at least two factors. First, Rechtschaffen's failure to get large reminiscence scores for his extraverts could be due to an insufficiently long rest pause; the rest given was only 1 min., which is considerably less than most investigators have used with the inverted alphabet printing task, e.g. Archer (1954) gave a 5 min. rest, and Kimble (1949) a 10 min. rest; and Eysenck obtained his correlation between extraversion and reminiscence with a 10 min. rest pause. It is likely that the 1 min. rest is insufficient for reactive inhibition to dissipate fully in extraverted subjects. Hence they would not show the full reminiscence effect obtained with a longer rest interval. Rechtschaffen argues that since it is postulated that extraverts generate reactive inhibition more quickly and dissipate it more slowly, conflicting predictions about reminiscence scores can be made from the Eysenck theory: either extraverts should dissipate more I_R because they have more and hence show more reminiscence; or because they dissipate I_R more slowly they should show less reminiscence; or the two effects might cancel each other out. This argument could easily explain Rechtschaffen's finding using his particular experimental conditions, but it cannot be accepted as a valid criticism of the Eysenck theory; for, once the time interval for reactive inhibition to dissipate in extraverts has been established by the standard method of plotting reminiscence scores with different rest intervals, the theory yields the quite unequivocal prediction that if this time interval is allowed extraverts will show more reminiscence.

A second factor which could explain Rechtschaffen's low correlation between extraversion and reminiscence arises from the part played by conditioned inhibition. Reminiscence scores are most satisfactorily obtained either very quickly, before sI_R has had time to develop, or after a number of trials, when sI_R has reached its asymptote. Rechtschaffen's choice of ten trials may be unsatisfactory from this point of view. Since reactive inhibition develops more quickly in extraverts, it is possible that by the eleventh trial they will have greater conditioned inhibition and this would obscure the recovery from reactive inhibition in the measurement of reminiscence. Since in the present study extraverts and introverts both work at the same

speed on the fifteenth trial, it appears that sI_R has reached its asymptote and the greater reminiscence of extraverts emerges.

The small correlation of neuroticism with duration of the spiral after-effect confirms previous reports that there is no association between the spiral after-effect and affective states (Price & Deabler, 1955). The correlation between neuroticism and time on trial 1 of the inverted alphabet printing task would be predicted on the basis of the Yerkes-Dodson law, i.e. the task is one in which the dominant habits are incorrect and are made more dominant by high drive, with the result that high drive impairs performance. Although the correlation falls short of significance, it is in the predicted direction and may be regarded as giving some support to the findings of Taylor & Rechtschaffen (1959) of a significant correlation between anxiety and time on inverted alphabet printing. On the other hand, the independence of neuroticism and reminiscence is inconsistent both with Eysenck's (1956) finding of a positive correlation between neuroticism and reminiscence on the pursuit rotor, and Kimble's (1950) finding of a correlation between drive and reminiscence in inverted alphabet printing. The discrepancy adds one more finding to the many reviewed by Jensen (1958) showing that experiments using questionnaire measures of drive give extremely conflicting results, probably determined in part by the degree to which drive is mobilized by the experimental situation.

The writer is indebted to Mr Ian Gordon of Exeter University for help in carrying out part of the experiment.

REFERENCES

- ARCHER, E. J. (1954). Post-rest performance in motor learning as a function of pre-rest degree of distribution of practice. *J. Exp. Psychol.* **47**, 47-51.
- EYSENCK, H. J. (1956). Reminiscence, drive and personality theory. *J. Abnorm. (Soc.) Psychol.* **53**, 328-33.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1959). Comments on a test of the personality-satiation-inhibition theory. *Psychol. Rep.* **5**, 395-69.
- EYSENCK, H. J., HOLLAND, H. & TROUTON, D. S. (1957). Drugs and personality. III. The effects of stimulant and depressant drugs on visual after-effects. *J. Ment. Sci.* **103**, 650-5.
- HAMILTON, V. (1959). Eysenck's theories of anxiety and hysteria—a methodological critique. *Brit. J. Psychol.* **50**, 48-63.
- JENSEN, A. R. (1958). Personality. *Ann. Rev. Psychol.* **9**, 295-322.
- KIMBLE, G. A. (1949). An experimental test of a two factor theory of inhibition. *J. Exp. Psychol.* **39**, 15-23.
- KIMBLE, G. A. (1950). Evidence for the role of motivation in determining the amount of reminiscence in pursuit rotor learning. *J. exp. Psychol.* **40**, 248-53.
- LYNN, R. (1959). Two personality characteristics related to educational attainment. *Brit. J. Educ. Psychol.* **29**, 213-6.
- PRICE, A. G. & DEABLER, H. L. (1955). Diagnosis of organicity by means of the spiral after-effect. *J. Cons. Psychol.* **19**, 299-302.
- RECHTSCHAFFEN, A. (1958). Neural satiation, reactive inhibition and introversion-extraversion. *J. Abnorm. Soc. Psychol.* **57**, 283-91.
- SPIVACK, G. & LEVINE, M. (1959). Spiral after-effect and measures of satiation in brain injured and normal subjects. *J. Personality*, **27**, 211-27.
- TAYLOR, J. A. & RECHTSCHAFFEN, A. (1959). Manifest anxiety and reversed alphabet printing. *J. Soc. Abnorm. Psychol.* **58**, 221-4.

(Manuscript received 26 May 1959)

ASSOCIATION BETWEEN MUSCULAR TENSION AND WORK OUTPUT*

By D. W. FORREST

Bedford College, University of London

E.M.G. electrodes were placed on the right forearm and chin of ten subjects while they performed simple addition. It was possible to record from one pair of electrodes the amount of muscular activity in the right forearm while the other pair indicated the moment of speaking and thus the speed of work.

The variations in muscular tension and in speed of work were followed during the course of addition and this intra-sum analysis revealed a different relationship from that found when larger units of work were taken. In the former case a marked tendency was found for increases in speed to be accompanied by low levels of tension, whereas when sums were taken as units no consistent relationship was apparent between the two variables within any one subject. This and other discrepancies are briefly discussed.

I. INTRODUCTION

There is a commonly held belief that intense mental work is accompanied by a set facial expression, furrowed brow and tense posture. The experimental literature appears on the whole to justify the opinion that a high level of bodily tension accompanies concentrated mental activity.

The earlier experimenters, who worked with the knee jerk as an index of muscular tension level, invariably obtained an increase in tension when work began (e.g. Lombard, 1887; Féré, 1889). More recent research has in general confirmed this finding. For example, Golla & Antonovitch (1929) used a simple apparatus for registering tension in the tibialis muscle and found that a steep rise in tension occurred as soon as subjects began addition. Jacobson (1929), who was the first to use electronic equipment, came to the conclusion that no mental activity occurred without concomitant measurable muscular tension, and Davis (1939) further elaborated these findings when he showed that particular patterns of muscular activity were associated with certain mental tasks.

Many workers have investigated the course of tension during the work period. A variety of tasks of different lengths has been employed and recordings have been made from many different loci. In spite of these differences, there is agreement that the steep initial rise in tension is sustained for at least a few seconds, but there is no such unanimity about the later course of tension. It has been found to increase, decrease, or to remain constant while work proceeds. Bills & Brown (1929), who obtained the last effect, used such an insensitive method of recording that any underlying trends may have been obscured. Other experiments fall into two groups: Golla & Antonovitch (1929), Freeman (1930), and Geldreich (1953) all of whom found that tension decreased until the end of the task, and Davis who obtained an increase in tension in two separate experiments (1937, 1939).

One striking difference between the work of Davis and of other experimenters lies in the length of the prescribed tasks. Davis's tasks took no longer than about

* This paper is based on a thesis presented for the Ph.D. degree of London University.

5 min., Freeman's were double this length, those of Golla & Antonovitch three times as long, and Geldreich's lasted for 55 min. As Davis in his 1937 paper presents curves of tension which decline towards the end of the work period, it seems a possible extrapolation to predict a continuous decrement if the task were to be continued. It should, however, be noted that Davis did not record from the same muscle groups as the other workers and that he employed discontinuous work. The counting and addition tasks used by the others were of a more cumulative and continuous nature.

A tentative conclusion can be derived from these experiments that at the beginning of mental work there is, in all muscle groups so far investigated, a marked rise in tension, which may be followed by a further increase to be succeeded in turn by a slow decline until the end of the task. The length and type of task probably influence the form of the tension curve and the locale of maximum change.

These fluctuations in tension which occur during the process of work might be thought to be associated with similar fluctuations in the rate of work, but little attempt has been made to relate the two. Since the time of Kraepelin (1902) it has been known that output at the beginning of the more repetitive type of mental task is usually negatively accelerated—the so-called 'starting spurt'—and that towards the end of the task an increase in speed occurs—the 'end spurt'.

Besides these terminal variations in speed, Voss (1899) and later Philpott (1932) have demonstrated periodic fluctuations in efficiency during the course of long-continued mental work.

There would seem to be a need to establish whether there is a relationship between such changes in work rates and alterations in muscular tension similar to those reported by other workers. The present experiment attempts to investigate this relationship by observing the variations in both tension level and speed of work during the process of mental addition.

II. THE PRESENT EXPERIMENT

Subjects. Ten members of a class in elementary psychology volunteered as subjects. They were all women between the ages of 18 and 22.

Apparatus. One channel of the electromyograph, which was of private design (see Forrest, 1956), was used to record muscular tension from the right forearm; bipolar sponge electrodes were attached half-way between the lateral humeral epicondyle and the styloid process of the ulna (Davis, 1953). The other channel recorded from above and below the point of the chin. This location gave an indication of the moment of speaking, and as subjects were required to add aloud it thus served as a record of the speed of work. The output of both channels was led to a Kelvin Hughes high-speed pen recorder which recorded on Teledeltos paper moving at 5 cm. per sec. A tape recorder was used so that the accuracy of the addition could be checked and any incidental observations made by the subject recorded.

The mental work consisted of addition sums typed vertically on white cards. Sums of three lengths were used: 6, 11, and 21 figures in a single column. The digits were chosen at random from Fisher & Yates' (1938) table of random numbers.

Design. Work was performed under three conditions:

Condition M (maximum)—addition at maximum speed;

Condition N (normal)—addition at an easy self-regulated pace;

Condition T (tension)—addition as in N, but with the subject pulling down on a spring balance with the right hand.

Under these conditions the possible relationships between the variables of work output and muscular tension were investigated. Condition M has been the usual approach of earlier investigators of mental work curves. It seemed possible that condition N might induce greater uniformity in work output, that is, that work would proceed with less marked spurts or blocks. Under condition T the muscular tension level of the subject would be deliberately disturbed. Thus conditions N and T represent attempts to alter respectively work rate and tension level from the values found in the most commonly used condition in order to examine more widely any relationship that may exist between the two variables in the M condition.

Three sums of each length were worked under each condition (27 sums in all). Ideally, sums of the three conditions and of the three lengths would have been randomized together, but on psychological grounds this was not considered practicable. The M condition required the subject to make a determined effort to work as quickly as possible and a rapid change of set from this type of sum to one where the subject had to work at her own speed did not seem possible. Therefore the sums of the M condition were worked in a block which preceded or followed the intermixed T and N conditions. Five subjects worked the M block first and five the intermixed T and N conditions.

In other words, it was presumed that changes in postural set (from N to T, or from T to N) could occur more readily than changes in mental effort; in any case an adequate check on the muscular condition of the subject could be obtained from the electromyogram, and it was a simple matter to wait until tension had reached a control level before allowing her to proceed with the next sum in the series.

Procedure. The subject lay on a couch inside a shielded booth. The addition sums were arranged in a simple tachistoscopic device above the subject's face at a comfortable distance for reading. She was told she would be shown figures in columns of various lengths which she was to add aloud and successively from top to bottom, ignoring tens and only adding units. (This adding procedure was adopted to secure uniformity of verbal response throughout the sum. If normal addition had been employed the later additions would have taken longer to say as they would have been polysyllabic numerals and this might have obscured variations in speed due to other causes.)

The electrodes were attached and trial recordings made until their placement was satisfactory. The spring balance was then demonstrated and the subject was asked to press on it with her right hand for half a minute. A reading was taken at the end of this period, and half this amount was employed during the T additions. (This procedure follows that of Stauffacher (1937) who found that the induction of this amount of tension led to the greatest improvement in the speed of mental work.)

The subject was then given a demonstration of the adding procedure and then worked at two practice sums. Instructions now differed depending on the order in which the subject was to be given the three types of work. If the M condition came first, instructions were as follows:

I want you to add these sums as quickly as you possibly can. Try to be accurate as everything you say will be recorded and checked later. But the main emphasis is on speed; go as fast as possible. I will time you over each sum. Now if you make a mistake, carry straight on. Don't go back to correct it. Is that clear?

For the N and T conditions instructions were:

I want you to add these sums at an easy speed. That is, disregard how long the sum is, just add at a speed which you could keep up for ten minutes at a stretch. I want you to be accurate as everything you say will be recorded and checked later. There is no rush and no emphasis on speed. Just add at an easy pleasant rate. If you make a mistake, please carry on. Don't go back to correct it. Is that clear?

In the case of the T condition the subject then pressed on the spring balance until the correct reading was obtained and this pressure was maintained during work.

After the instructions had been understood the subject was requested to relax until no spikes of tension could be seen on the C.R.T., whereupon the first sum was given. A delay of from 2 to 10 min. occurred between sums until the control level of tension had been re-attained; a rest of 10 min. was taken between the M and N and T blocks.

The detailed instructions were only given before the first sum of each condition was added. For later sums in the series abbreviated instructions were given. For example, if an M-type sum was to be added the experimenter would say, 'Once again as quickly as you can'.

III. RESULTS

Quantification of voice record. The first obvious spike of more than 2 mm. in height of the cluster of spikes in the E.M.G. record that occurred when the subject spoke was taken as the moment of verbalization. Distances between successive initial spikes were measured to the nearest mm. and converted into sec. (1 mm. = 0.02 sec.). Thus there were five time intervals in the case of the six-figure sums, ten in the eleven-figure sums, and twenty in the twenty-one-figure sums.

Quantification of tension record. A pair of dividers set to 2.5 mm. ($\frac{1}{20}$ sec.) was used to step off small divisions from one verbalization to the next. These $\frac{1}{20}$ sec. divisions were then numbered. One-quarter of these divisions were now chosen at random and all the spikes within the chosen divisions measured with a transparent ruler to the nearest $\frac{1}{2}$ mm.

The mean of these measurements was calculated for each division and the mean of the several divisional means gave a figure representative of the interval as a whole. Thus a tension reading was obtained for the intervals between each spoken digit of each sum, and from these a mean for the whole sum could also be derived.

Treatment of results. We will not present here more than an outline of the changes in rate of work and tension level brought about by the three conditions under which work took place. These facts, which have been reported in detail elsewhere (Forrest, 1958), can be briefly summarized as follows. Under a set to work quickly (condition M) tension levels were higher than under condition N, and this association between high tension and high output was also found in condition T where the induced tension led to a more rapid rate of work.

In the present analysis we are concerned with the relationship between tension and work output during the process of addition and with the effects of the three conditions on any such relationship. The examination can take place at two levels: (1) As we know the total time taken to add a sum and the total amount of tension exerted during it, correlations can be run between these totals either (a) within a subject, or (b) between subjects. (2) On a different (intra-sum) level, as we have a record of the

time taken over each addend of a sum and the accompanying tension level, it is possible to correlate these fluctuations in speed and in tension which occurred during work.

(1a) With regard to the first possibility, when the individual records were examined the correlations could be seen to vary enormously from one subject to the next. The variation was from +1.0 to -1.0, approximately half the correlations being positive and half negative. With more subjects it might have been worth while treating in separate groups those who had positive and those who had negative correlations. The difference between these groups is possibly a key to some of the seeming contradictions which will emerge later in this analysis.

(1b) When the association between subjects was examined a more consistent picture emerged. Table 1 reveals a tendency for time and tension to be positively correlated; that is, those subjects who have added more slowly tend to be more tense. (The average correlations in Table 1 have been obtained from the correlations for each subject by the use of Fisher's z transformation.)

Table 1. *Correlations between total time and total tension between subjects*

Condition	Short	Medium	Long	Average
T	0.39	0.15	0.06	0.21
N	0.45	0.71	0.32	0.52
M	0.12	0.26	0.22	0.20
Average	0.33	0.38	0.21	0.32

(2) To pass to the intra-sum level: in the short sums product moment correlations were computed between two sets of five figures, the five readings for tension and the five readings for time, in the sums of medium length between two sets of ten figures, and in the long sums between two sets of twenty figures. Each set of figures was an average taken over the three sums of each length added by each subject under each condition. As the sums were randomized, these correlations do not represent the relationship between the two variables over particular pairs of addends, but rather the general relationship between time and tension since the sum was begun.

Nine correlations were obtained for each subject in this way. Of the total of ninety correlations, eighty-nine were positive and extremely high. Sixty-four of them reached the 5% level of significance and these tended to cluster in the long sums. However, it is not possible to interpret these significance levels literally as the coefficients were calculated from readings that were not independent of one another. That is, the speed of work and the tension exerted at one stage in the addition of a sum were likely to influence the speed and tension at a later stage.

These data were transformed into z scores and an analysis of variance carried out. The results of the analysis are given in Table 2 from which it can be seen that only two sources of variance are significant, that due to lengths (at the 0.01 level) and that due to conditions (at the 0.05 level).

The mean z scores are given in Table 3 and reveal that the medium and long sums are alike and differ from the short sums at the 0.01 level. Similarly, the T and N conditions resemble one another but differ from the M condition, the difference again being significant at the 0.01 level. It can be concluded that the relationship between

time and tension is most marked in the short sums and in those of the M condition, and this relationship seems to be independent of differences between subjects.

Table 4 gives average correlation coefficients obtained by transformation of Table 3. The overall positive correlation is very high (+0.72). Thus, there is a marked tendency for high tension to be associated with a slow rate of work when these variables are examined at an intra-sum level.

Table 2. *Analysis of variance of 'z' scores*

Source	D.F.	Sums of squares	F
Lengths	2	2.9166	15.54**
Conditions	2	1.7392	9.27*
Lengths $\times$ conditions	4	0.1876	0.012
Subjects	9	1.0773	0.347
Subjects $\times$ lengths	18	3.0996	0.41
Subjects $\times$ conditions	18	4.869	0.644
Subjects $\times$ lengths $\times$ conditions	36	7.5564	0.487
Interaction	72	15.505	—
Total	89		

* $P < 0.05$; ** $P < 0.01$.

Table 3. *Mean 'z' scores*

Condition	Short	Medium	Long	Average
T	1.049	0.743	0.620	0.804
N	1.007	0.753	0.666	0.809
M	1.358	1.133	0.813	1.101
Average	1.138	0.876	0.700	0.905
		0.788		

Table 4. *Average correlation coefficients between time and tension within sums*

Condition	Short	Medium	Long	Average
T	0.78	0.63	0.55	0.66
N	0.77	0.64	0.58	0.67
M	0.88	0.81	0.67	0.80
Average	0.81	0.71	0.60	0.72

IV. DISCUSSION

To recapitulate: an analysis of the relationship between performance and tension was attempted at two levels,

- (1) between the total time taken over a sum and the total tension recorded during it (a) within individual subjects, and (b) between subjects, and
- (2) between the time taken over each addend of a sum and the tension accompanying it.

In the case of (1a) no consistent relationship was apparent, in (1b) a low but consistently positive correlation occurred, and in (2) a high positive correlation was found which was most marked in the short sums and in those which were added under the set to work quickly.

Although the overall picture seems fairly consistent, (1a) and (1b) are difficult to

reconcile. The fact that total time and total tension correlate when subjects are compared seems to point to differences in habitual tension and associated differences in speed of work. Those who added slowly had a higher tension level. But the lack of correlation when the totals for the sums of any one subject are compared seems to imply that 'chance' variations were superimposed on the habitual tension level and that these did not affect speed.

The absence of correlation between time and tension in the totals of the sums of any one subject is all the odder in view of the very high correlation when addends are considered. During the process of work high tension was closely linked with slow speed and yet when the sum is taken as a whole no association is found. It is as though subjects began a sum at a random level of tension (within their habitual limits, which would be higher for the slow worker). If the sum were a long one, subjects behaved alike in tending to be tenser. But irrespective of the type and length of sum, tension fluctuated during work in an inverse relationship to speed of work. When the subject was asked to work quickly, or was given a short sum, this inverse relationship was very clearly defined.

Further difficulties arise when the general role of tension is considered in its relation to efficiency. For, as we saw earlier, induced physical tension (condition T) led to more rapid work, and also higher tension levels were recorded under the M condition of faster work. The former finding is not, however, directly comparable with the apparently similar enhancement obtained by earlier workers, notably Stauffacher (1937). In other reported work the effects of tension have been superimposed on maximal mental effort. In the present experiment tension was induced in subjects who had been requested to work at an easy steady rate. The changes that occurred in their speed of work may be due to either (1) a direct physiological effect brought about by the induced tension, or (2) an indirect effect due to failure to retain the set to work slowly. This failure might be due to a misunderstanding of the instructions, especially in the shortened form in which they were given after the first trial under the T condition, or to a real difficulty in working slowly while tensing the forearm. The replies of two subjects who were questioned some time after the experiment support the idea that a change of set occurred. Both said that they had to keep reminding themselves not to go too fast while pulling on the spring balance. Unfortunately further retrospective evidence was not available.

Of course, results obtained from experiments in which tension is induced are not strictly comparable with those obtained by monitoring a subject's 'natural' tension during a task. The former type of experiment produces very large disturbances in any 'natural' pattern of tension which may exist prior to the induction. Furthermore, an increment to the tension in the forearm may lead to a lessening of the tension elsewhere. Thus, the enhanced efficiency under the T condition in the present experiment, whether brought about 'directly' by the local increase in tension or by a change in set, need not present difficulties for any theory wishing to relate a high level of tension to a slow rate of work.

It is otherwise, however, with the results obtained in the M condition. Here subjects became more tense under a set to work quickly. The tension in this case may have been a by-product of the instructions and in no way a necessary condition for the rapid work. But the contradiction needs to be investigated further.

It will be remembered that the slower workers tended to have higher tension levels. This finding is in agreement with much earlier work showing that tense people work slowly. For example, anxiety neurotics, who are not renowned for their efficiency, have higher tension levels than normals (Shagass & Malmo, 1954; Sainsbury & Gibson, 1954); their tension appears neither to aid the solution of every-day problems nor to improve performance at psychological tests. The work of Jacobson (1929) and Schultz (1956) on the therapeutic value of relaxation supports this contention, and an experimental demonstration is to be found in Duffy's research (1932) which appears to show that highly tense children were not the most efficient performers.

The present results from the intra-sum analysis (although not a comparison between subjects) lead to the same conclusion, the higher the tension the lower the output. It seems likely in view of the high correlation at this level that the time and tension curves parallel one another throughout, and that the deceleration in rate of work from the beginning of a task noted by the early German experimenters is associated with the increase in tension reported by other workers, and that the acceleration towards the end is similarly associated with a decrease in tension such as Geldreich (1953) has reported. This would imply that starting and end spurts can occur in tasks of very short duration; the present writer has some evidence that this is so (Forrest, 1960).

This finer analysis of what occurs within a unit of work appears to be profitable and is in line with the modern analysis of skilled performance and its disturbance where total output is no longer regarded as sufficient. But confirmation is needed before these rapid oscillations in speed of work and in level of muscular tension can be accepted as being more than artifactual. Previous work has not, as far as I know, been concerned with output on such a fine time scale nor over tasks of such short duration; it would be unwise to extrapolate to longer tasks even of a similar nature.

REFERENCES

- BILLS, A. G. & BROWN, C. (1929). The quantitative set. *J. Exp. Psychol.* **12**, 301-323.
- DAVIS, J. F. (1953). *Manual of Surface Electromyography*. Laboratory for Psychological Studies, McGill University, Montreal.
- DAVIS, R. C. (1937). The relation of certain muscle action potentials to 'mental work'. *Ind. Univ. Publ. Sci. Ser.* no. 5.
- DAVIS, R. C. (1939). Patterns of muscular tension during 'mental work' and their constancy. *J. Exp. Psychol.* **24**, 451-65.
- DUFFY, E. (1932). The relation between muscular tension and quality of performance. *Amer. J. Psychol.* **44**, 535-46.
- FÉRÉ, C. S. (1889). In BILLS, A. G. (1927). The influence of muscular tension on the efficiency of mental work. *Amer. J. Psychol.* **38**, 227-51.
- FORREST, D. W. (1956). *A study of electromyographic changes associated with mental work*. Unpublished Ph.D. thesis, University of London.
- FORREST, D. W. (1958). Influence of length of task on rate of work and level of muscular tension. *Occup. Psychol.* **32**, 253-7.
- FORREST, D. W. (1960). The starting spurt in mental addition. (In preparation.)
- FREEMAN, G. L. (1930). Changes in tonus during completed and interrupted mental work. *J. Gen. Psychol.* **4**, 309-33.
- GELDREICH, E. W. (1953). Some physiological concomitants of mental work. *Psychol. Monogr.* No. 358.
- GOLLA, F. L. & ANTONOVITCH, S. (1929). The relation of muscular tonus and the patellar reflex to mental work. *J. Ment. Sci.* **75**, 234-41.

- JACOBSON, E. (1929). *Progressive Relaxation*. Chicago: University Press.
- KRAEPELIN, E. (1902). Die Arbeitscurve. *Philos. Stud.* **19**, 459-507.
- LOMBARD, W. R. (1887). The variations in the normal knee jerk. *Amer. J. Psychol.* **1**, 5-71.
- PHILPOTT, S. J. F. (1932). Fluctuations in human output. *Brit. J. Psychol. Monog. Suppl.* No. 17.
- SAINSBURY, P. & GIBSON, J. G. (1954). Symptoms of anxiety and tension and the accompanying physiological changes in the muscular system. *J. Neurol. Neurosurg. Psychiat.* **17**, 216-24.
- SCHULTZ, J. H. (1956). *Das Autogene Training (konzentrierte Selbstentspannung)*. G. Thieme, Stuttgart.
- SHAGASS, C. & MALMO, R. B. (1954). Psychodynamic themes and localized muscular tension during psychotherapy. *Psychosomat. Med.* **16**, 295-314.
- STAUFFACHER, J. C. (1937). The effect of induced muscular tension upon various phases of the learning process. *J. Exp. Psychol.* **21**, 26-46.
- VOSS, G. VON (1899). Ueber die Schwankungen der geistigen Arbeitsleistung. *Psychol. Arbeiten.* **2**, 399-449.

(Manuscript received 17 November 1959)

THE EFFECT OF DISTRACTION ON TIME GAP IN THE ESTIMATION OF SHORT DURATION

By R. G. CHATTERJEA

Psychology Laboratory, Calcutta University

An experimental study was performed on the nature of time gap (i.e. the time before the subject begins his estimation) in relation to estimation of short duration of different magnitudes under an ambient condition of noise. The estimates of all subjects at the different stimulus durations were found to be overestimations. The magnitude of the time gap showed a greater correlation with stimulus duration than that found in a previous study, without noise.

I. INTRODUCTION

In a previous study (Chatterjea, 1957), it was shown that a very small amount of time elapses between the presentation of a stimulus time and subsequent estimation of the same by the subject. This intervening time has been termed 'time gap' to differentiate the phenomenon from 'pause'. The term 'pause' is referred to a situation where the intervening period is controlled, whereas the term 'time gap' as envisaged here is that time which an individual apparently feels he needs before estimating a time duration and is not, of course, subject to control by the experimenter. The present paper deals with an experiment on the effect of auditory background on apparent duration.

II. OUTLINE OF PREVIOUS RESEARCH

Experimentation on the estimation of time appears to have begun with Czermak (Titchener, 1905), though elaborate studies on the various problems are of much later origin. But the psychophysics of time perception are primarily concerned more with the study of intervals rather than that of duration. Notable contributions in the field of interval estimation by Woodrow and others have been reviewed by Boring (1942) and Woodrow (1951). Studies on the estimation of duration, on the other hand, are fewer in number and have been carried out by different investigators working in different places with entirely different procedures.

The effect of four combinations of ambient conditions of light, dark, quiet and noise is reported by Hirsh *et al.* (1956), where durations of 1, 2, 4, 8 and 16 sec. were employed. These investigators, however, left the intervening period between the cessation of the stimulus time and the judgement uncontrolled and unmeasured. The work of Stott (1935) with continuous tones of varied lengths is mainly focused on the evaluation of the time order errors. Studies by Cohen *et al.* (1954) and Sturt (1923) differ from the present writer's (Chatterjea, 1956, 1957, 1958) primarily in procedural techniques.

III. PLAN OF PRESENT INVESTIGATION

(i) *Subjects.* The subjects were students of senior postgraduate classes in psychology. They were seven men and three women selected at random from the class register. The subjects were cheerful and co-operative throughout the experimental sitting.

(ii) *Materials.* The materials used were all standard instruments employed in time-perception experiments, viz. the Leipzig model of the Time Sense Apparatus, the Hipp Chronoscope, the Ludwig Baltzar model Universal Kymograph, tuning fork time marker, electric bell, etc. The Time Sense Apparatus (TSA) was electrically driven by a variable speed motor.

(iii) *Experimental design.* The design performed three functions, viz. (a) accurate presentation of the stimulus times, (b) efficient recording of the judgements, and (c) translating the tracings into units of seconds. By suitable adjustments it was so arranged that the difference between the two arms of the TSA produced the desired duration. The stimulus time appeared in the form of a continuous light. The duration of the light was, however, checked on a Hipp chronoscope. The second function was secured by the use of kymograph paper of sufficient length to accommodate at least five judgements in the course of one complete rotation of the drum. The stimulus times as well as the estimations were recorded on the drum paper, the time gap being recorded automatically in between these two. A 100 cps tuning fork having electromagnetic arrangements was used to serve the purpose of a time marker.

An electric bell was placed close to the subject. It was so arranged that the bell started ringing as the light denoting the stimulus time went off, and the bell stopped when the subject had finished giving his or her estimation. There was, however, no extraneous sound during the presentation of the stimulus time. At the start of the experimental sitting each subject was asked not to adopt any means like counting, making rhythmical movements, etc. The following was then read out: '*The bulb in front of you will glow for a certain length of time. You are to estimate the time the light was on. You press the key and keep it pressing so long as you think the bulb was glowing. You will hear a bell ringing while you are assessing the time. Do not bother for this sound but try to be as accurate as possible in your judgements.*'

There were ten stimulus times, of 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.00 sec., each bounded by a continuous light of low intensity. Each stimulus time was presented ten times, the subject giving ten judgements. When estimation of one stimulus time by one subject had been obtained, the next stimulus duration was presented and the sequence continued till the last stimulus time.

IV. RESULTS AND DISCUSSION

The method used in this study is the method of reproduction, and procedures for the analysis of the data are the same as indicated by Woodrow (1951). The mean estimates of each stimulus time are shown in Table 1, along with their constant errors (C.E.) and standard deviations (S.D.).

The second section of Table 1 presents the mean time gaps at each stimulus time.

The intercorrelations between means of estimates and means of time gap are shown in the final section of Table 1.

In experiments with the method of reproduction there are to be found objective and subjective errors. The former include errors due to adjustments of the arms of the TSA, inaccuracy in the recording instrument, fluctuations of current, etc. The variable influences of these have a bearing on the precision of the stimulus times, but this was eliminated by actual counting from the drum paper.

The other kind of error, viz. the subjective error, was sought to be reduced by frequent repetition of the instructions. Influences due to inattention, boredom, etc., on the part of the subjects were also ruled out, since the subject was required to participate for only a short time each day.

The effect of a third type of error, the quotidian variability, was similarly eliminated, since the subject gave all his judgments at any stimulus time at one sitting only.

The mean estimate of each subject rises in magnitude along with the rise in the magnitude of stimulus times. Similar trends had been obtained in a previous study

Table 1. *Estimates, time gap and correlations between these for different stimulus durations*

Stimulus duration ...	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
Estimates										
Mean	0.160	0.229	0.340	0.430	0.508	0.613	0.746	0.856	0.949	1.059
S.D.	0.017	0.024	0.013	0.011	0.011	0.019	0.038	0.025	0.036	0.046
C.E.	+0.060	+0.029	+0.040	+0.030	+0.008	+0.013	+0.046	+0.056	+0.049	+0.059
Time gap										
Mean	0.317	0.362	0.406	0.436	0.484	0.514	0.552	0.577	0.617	0.631
S.D.	0.042	0.036	0.034	0.031	0.038	0.030	0.042	0.060	0.052	0.033
Correlations between estimate and time gap										
<i>r</i>	+0.556	+0.224	+0.017	-0.112	+0.085	+0.339	+0.664	+0.842	+0.737	+0.735

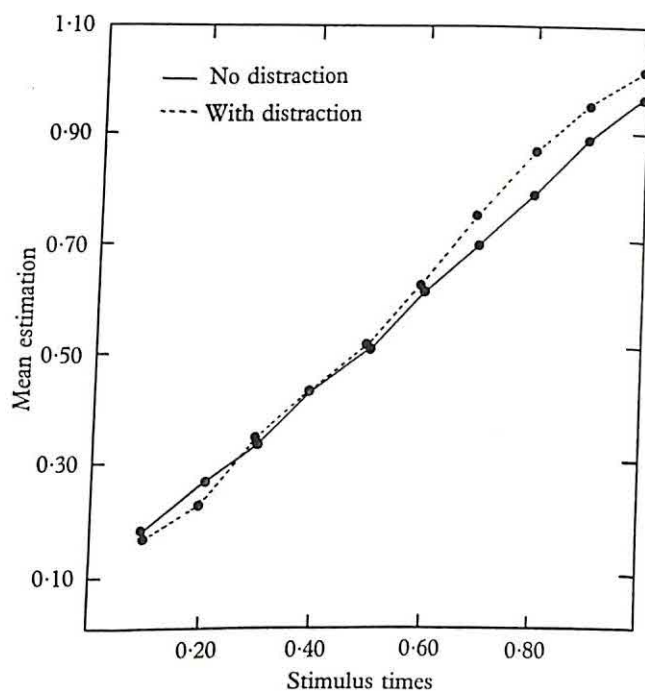


Fig. 1. Mean estimations under distracting and non-distracting conditions.

(Chatterjea, 1957). The rise in the estimate values for the group is similarly conspicuous, though this rise is not regular. It appears that the dependence of estimations on stimulus duration is close and the regression relation is approximately linear, as seen in Fig. 1.

Whereas previous studies have shown both underestimation and overestimation of the stimulus times, all the estimates in the present experiment save a few are positive. This overestimating bias is common to all the subjects irrespective of stimulus presentations. In view of this the concept of indifference duration as applied here is the stimulus duration at which the mean value of the mean error of a subject is very nearly zero, in contrast to the usual definition of the indifference point, which supposes it to lie between underestimates and overestimates, but it is found that there is no one indifference point that can fruitfully be applied for all the cases. While for S 1 the point lies between stimulus durations 0.50–0.60 the indifference point recedes back to 0.40–0.50 sec. for S 7 and even rises to 0.60–0.70 sec. durations for another subject. One subject (S 6) has shown zero error. The best we can do is to consider the mean value of the mean errors as locating the indifference point. The group C.E. as shown in Table 1 shows that the indifference may lie in the stimulus range 0.50–0.60 sec. Similar results appear to have been found by Woodrow (1951) and Chatterjea (1957).

A striking feature in the magnitude of the C.E.'s is that constant error is relatively greater at both ends of the stimulus series, while in the middle ranges it is comparatively small. The group C.E.'s similarly present a trend that is in conformity with the writer's previous findings.

The magnitudes of the time gap rises along with stimulus duration rather regularly and the relation between the two is nearly linear (Fig. 2). But there are certain discrepancies in the individual time-gap values in that some individuals show shorter time gaps than others.

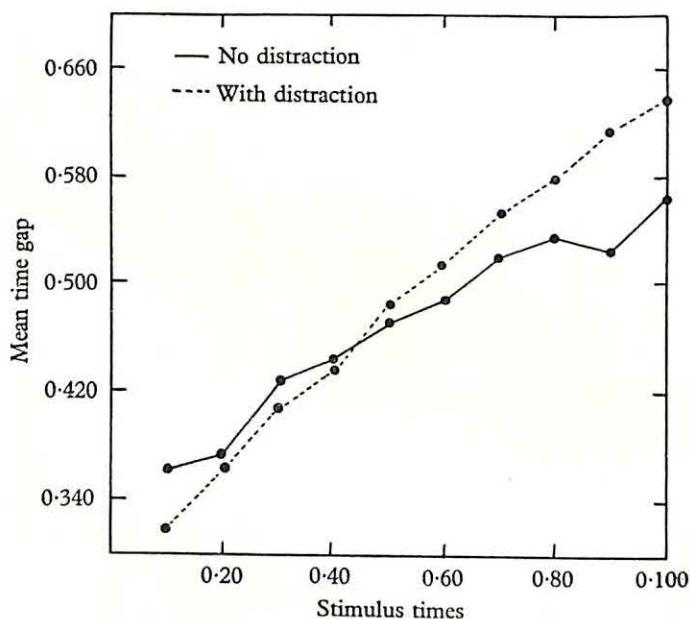


Fig. 2. Mean time gaps under distracting and non-distracting conditions

That the trend in the rise of time gap values is somewhat irregular in the noiseless condition will be evident from the above figure. In yet another experiment (Chatterjea, 1959) where the stimulus durations were presented through the auditory modality, it

was found that the nature of time gap was rather erratic, showing a much wider divergence from the present experiment. The inference drawn here is that in noisy conditions the time gap increases steadily along with stimulus time, and that it is greater than when no distraction exists. Distraction therefore appears to influence judgement in time estimation and consequently increases magnitude of time gap. This generalization is, however, only partially correct in that the increase in time gap values are mainly at the upper end of the stimulus series. A glance at the correlation table will substantiate this statement. The coefficients are all positive save one, but only four are significant. The interrelation between estimation and time gap, too, will be found to be most clearly marked at the upper end of the stimulus series.

V. CONCLUSION

In an attempt to show the effect of a distracting situation in the estimation of short time durations the following results have been obtained:

- (1) All the stimulus times are overestimated by all the subjects.
- (2) The rise in time gap along with increase in the stimulus times is more conspicuous than those obtained without any distracting situation.

REFERENCES

- BORING, E. G. (1942). *Sensation and Perception in the History of Experimental Psychology*. New York: Appleton Century.
- CHATTERJEA, R. G. (1956). The estimation of temporal intervals in different sense modalities—an experimental study. *Indian J. Psychol.* **31**, 121-4.
- CHATTERJEA, R. G. (1957). An experimental study on the nature of time gap phenomenon in the estimation of short duration in two sense modalities. Doctoral Thesis, Calcutta University.
- CHATTERJEA, R. G. (1958). Estimation of short temporal duration in two sense modalities by the method of reproduction. *Indian J. Psychol.* **33**, 21-4.
- CHATTERJEA, R. G. (1959). Time gap in the estimation of short duration with auditory presentation. *Aus. J. Psychol.* **11**, 182-90.
- COHEN, J., HANSEL, C. E. M., and SYLVESTER, J. (1954). An experimental study of comparative judgements on time. *Brit. J. Psychol.* **45**, 108-14.
- HIRSH, I. J., BILGER, R. C. and DEATHERAGG, B. H. (1956). The effect of auditory and visual background on apparent duration. *Amer. J. Psychol.* **69**, 561-74.
- STOTT, L. H. (1935). Time order errors in the discrimination of short tonal duration. *J. Exp. Psychol.* **18**, 741-66.
- STURT, M. (1923). Experiments on the estimate of duration. *Brit. J. Psychol.* **13**, 382-8.
- TITCHENER, E. B. (1905). *Experimental Psychology*, vol. II, part 2. New York: MacMillan.
- WOODROW, H. (1951). Time perception. In *Handbook of Experimental Psychology*, chap. 32. Ed. by S. S. Stevens.

(MS. received 9 December 1958—Revised 17 May 1960)

THE EFFECT OF ILLUMINATION CONDITIONS ON THE EFFECTIVENESS OF LIGHT-ONSET AS A REINFORCER: A TEST OF THE BEVAN-ADAMSON REINFORCEMENT THEORY*

By HARRY M. B. HURWITZ

Birkbeck College, University of London

An experiment was performed to test a deduction from the Bevan-Adamson theory of reinforcement, namely, that the reinforcing effectiveness of a stimulus (light-onset) is dependent on the adaptation to background stimuli and an internal referent. Two groups of hooded rats (five in each group) were given seven half-hour operant sessions in a dark and an illuminated box, respectively. A measure of the frequency of bar pressing and exploration of a copper screen was taken. The groups were then given fourteen sessions in a dark box, bar pressing resulting in light-onset. The results showed that the bar pressing rate for the dark operant-tested subjects was significantly higher than for subjects operant-tested in an illuminated box. These results appear to be in accordance with the Bevan-Adamson theory.

Despite demonstrations of the efficacy of light-onset and, to a lesser extent, light-offset as a reinforcer for the bar-pressing response in rats (Girdner, 1953; Marx, Henderson & Roberts, 1955; Hurwitz, 1956), it remains problematic why this stimulus possesses reinforcing properties. An early suggestion (Marx *et al.* 1955) was that stimulus change, or novelty, may serve as a positive reinforcer in its own right rather than being 'due to some kind of unrecognized adventitious establishment of secondary reinforcing powers'. However, later studies explored the possibility that light-onset had in fact acquired secondary reinforcing powers. Thus Henderson (1957) concluded in favour of the secondary reinforcement hypothesis; Roberts, Marx & Collier (1958) found no support for it; whereas Hurwitz & Appel (1960) showed that the degree of reinforcing power was temporarily affected by associating light-onset with food presentation although light-onset reinforces bar-pressing responses independently of such previous association.

More recently, Stewart (1959) has put forward a theory to account for the results of a factorial experiment in which four different light intensities and six different reinforcement schedules served as independent variables. The results showed that as the ratio of reinforcement is increased, high response rates occur only with high light intensities. More specifically, given a low ratio of unreinforced to reinforced responses, the response rate increased for all four light intensities used. However, as the ratio is increased, the response rate increases at the higher light intensities and decreases at the low intensities. Stewart reasoned that the significant variable determining such interaction between light intensity and the reinforcement schedule may be the 'amount of stimulation per unit time brought about by the response, where "amount" is a function of both intensity of stimulation and the time interval between successive reinforcements'. Her hypothesis finds collateral support in a number of studies, for example, Cockrell (1952), Guttman (1953), Collier & Siskel (1959) and Kaplan (1957).

* The project was supported by a small grant from the Central Research Fund, University of London.

Stewart's thesis is closely related to a quantitative theory of reinforcement formulated by Bevan & Adamson (1960). Since the experiment reported in this paper was designed to test an implication of their theory, a brief outline is appropriate:

The presentation of a stimulus, S , gives rise to stimulation processes within the organism which becomes differentiated into two components, s_p (primary stimulation) and s_b (background stimulation). Over time, the stimulation is averaged so as to form an internal standard or referent ($\bar{s}$) which is a *product* of two components, viz.

$$\frac{(s_{b1} + s_{b2} \dots s_{bn})^x}{N} \quad (1)$$

and
$$\frac{(s_{p1} + s_{p2} \dots s_{pn})^y}{N} \quad (2)$$

where (1) is an expression to represent the average intensity of a series of background stimulus inputs, and (2) is the average intensity of a series of primary stimulus inputs, and where the exponents x and y indicate the relative contributions of s_b and s_p to the norm, $\bar{s}$. In order that a stimulus behaves like a reinforcer, a further condition has to be satisfied. The difference between s_p and $\bar{s}$ must be positive and the tension level of the organism (T) to which both physiological and stimulus input variables contribute, should be placed at the optimum region of the tension-performance curve. The effectiveness of reinforcement is then said to be a function of the respective values of ($s_p - \bar{s}$) and T .

One implication of the theory is that increasing the magnitude (intensity) of the reinforcer under given stimulus background conditions results in progressively more rapid alterations in performance. Further, the theory suggests that the reinforcement value of a stimulus is a function of the degree of adaptation of the organism to the background stimulus conditions, or experimental situation (s_b). If the average stimulation of s_b is low, a particular intensity of s_p will be more effective as a reinforcer than if s_b is high. This aspect of the theory can be tested in the following manner, using light-onset as the reinforcing stimulus for the bar-pressing response: the experimental subject, a rat, is placed either into a dark or an illuminated apparatus. Following a series of such exposures (adaptation trials), the animal is tested in a dark apparatus, light onset being made contingent on bar pressing. Under these conditions the theory predicts greater bar pressing activity for an animal adapted to the dark apparatus than for one adapted to an illuminated apparatus.

METHOD

Subjects. The subjects were ten experimental naïve male hooded rats, 90 days of age at the beginning of the experiment. They were drawn from two litters from the closed colony of hooded rats maintained by the animal laboratory of the Department of Psychology. Animals were randomly assigned to either of two groups.

Apparatus. The apparatus consisted of two identical boxes measuring 9 in. long $\times$ 12 in. wide $\times$ 14 in. high, which were placed into two separate sound-proofed cabinets. The boxes were painted a dull black and were fitted with (a) a bar, (b) a window, and (c) a metal floor. The bar was retractable (solenoid operated) and extended $\frac{3}{4}$ in. into the box. It was oblong shaped, 1 in. long and $\frac{1}{2}$ in. wide and responded to a pressure of 7–10 g. On the opposite wall to the bar was cut a $4 \times 3\frac{3}{4}$ in. window covered by an opaque glass and a fine copper gauze screen. The screen could be illuminated from the rear by a 6 V. lamp, giving approximately 0.50 ft.c.p. The

solid metal floor and the copper gauze screen covering the window constituted the terminals of an electronic contact relay, so that when the animal touched the screen, this was recorded on a counter.

Each box was permanently housed in a sound-proofed cabinet, consisting of an inner and an outer chamber. The inner chamber was made from $\frac{1}{2}$ in. thick celotex and measured 3 ft. long, 1½ ft wide and 1 ft. high. It was suspended in the outer cabinet by chair webbing. The outer cabinet was made from 2 in. thick compressed straw and measured 4 ft. long, 2 ft. 4 in. wide. The space between the two cabinets was filled with slag wool. The control equipment consisted of two units: (a) the light reinforcement control unit and recorder; (b) the electronic contact relay. The experiment was conducted in a darkened room adjoining the main animal laboratory.

Procedure. One week before the first experimental session, the animals were placed into an automatic feeding battery stationed in the main laboratory which made food available for 2 hr. daily at a time set aside for experimentation. This feeding schedule was subsequently maintained for the duration of the study.

Phase I. Each subject was given a daily $\frac{1}{2}$ hr. session in the apparatus approximately 1 hr. before feeding time, for 7 days. Group I was placed into a dark box; group II into an illuminated box. The frequency of bar pressing and window touching was recorded.

Phase II. Groups I and II were given daily $\frac{1}{2}$ hr. sessions for 14 days: whenever the bar was pressed the window was illuminated for the duration of the press. When the bar was released the box was once more plunged into darkness.

RESULTS

The results are summarized in Tables 1 and 2. Briefly, the operant level for the bar appears to be affected by the prevailing conditions of illumination and has a highly significant bearing on the response rate during light-onset reinforcement sessions.

Table 1. Comparison of bar response frequencies for groups' operant level tested in darkness (group A) and illumination (group B) and their performance during light-onset reinforcement

Conditions	Bar response frequency	Significance test
Phase I		
Group A, box dark	38.8	$t = 2.0, P = 0.10$
Group B, box lit	19.1	
Phase II		
Box dark: group A	142.1	$t = 6.0, P < 0.01$
Bar press results in light-onset: group B	47.1	
Group A: phase I vs. II	—	$6 = 7.1, P < 0.01$
Group B: phase I vs. II	—	$6 = 4.5, P < 0.01$

(a) Conditions of illumination and the operant level for bar and window responses. During phase I, group A was placed into a dark box whereas group B was placed into an illuminated box with the light source positioned behind the window. The results show that the frequency of lever and of window responses for the two groups do not differ more than would be expected by chance. (The difference between the two groups for bar pressing is significant at 0.10 only.) In this respect the results are in agreement with those reported by Robinson (1959).

(b) The effect of the independent variable during light-onset reinforcement sessions on the bar and window response output.

In phase II both groups were placed into a dark box and bar pressing was made to result in light-onset. Tables 1 and 2 show that

(a) Group A increased its bar-pressing rate from 36.8 to 142.1 and group B from 19.1 to 47.1. A *t*-test for correlated means reveals these differences to be significant beyond the 0.01 level.

(b) Group B decreased its frequency of window responses to a significant degree, the mean for phase I being higher than for phase II (34.3 *vs.* 12.0, respectively).

(c) The difference in bar-pressing rate between groups A and B during phase II is highly significant ($t = 6.0$, $P < 0.01$).

(d) The difference in window responses was not significant between the groups ($t = 1.91$, $P = 0.1$).

(e) The difference between the bar pressing in phase I and in phase II is greater for group A than for group B ($t = 5.1$, $P = 0.01$).

(f) The difference between window responses in Phases I and II is smaller for group B than for group A ($t = 2.4$, $P = 0.05$).

Table 2. *Comparison of window response frequencies for groups' operant level tested in darkness (group A) and illumination (group B) and their performance during light-onset reinforcement.*

Conditions	Window response frequency	Significance test
Phase I		
Group A: box dark	29.3	$t = 0.6$, not sig.
Group B: box lit	34.3	
Phase II		
Box dark: Group A	20.8	$t = 1.91$, not sig.
Bar press results in light-onset: group B	12.0	
Group A: Phase I <i>vs.</i> II	—	$t = 1.2$, not sig.
Group B: Phase I <i>vs.</i> II	—	$t = 6.5$, $P < 0.01$

DISCUSSION

Various attempts have been made to give the results of the numerous light-onset and light-offset studies a theoretical setting.

Early investigators (e.g. Kish, 1955; Marx *et al.* 1955) reasoned that the results were in accordance with the hypothesis that 'a perceptible environmental change, which is unrelated to such need states as hunger or thirst, will reinforce any response that it follows' (Kish) and that this supports an interpretation in terms of novelty as a factor in reinforcement.

The hypothesis of novelty, however, is too weak to mediate the prediction of results so far obtained. The novelty hypothesis demands an equal effect for both light onset and offset. Yet Hurwitz (1956), Roberts *et al.* (1959) and Robinson (1959) have reported that light-offset provides a weaker effect than light-onset. Secondly, the novelty hypothesis requires that a systematic decline in the response rate occurs both within sessions and from session to session. This does not seem to be the case. Under some conditions there is an actual increase in response rate per session (Robinson, 1959; Hurwitz, 1956) although in general the highest response rates are recorded on the first light-onset session, after which the rate levels

off to a value between the operant rate and the first reinforcement session (Stewart, 1959; Appel & Hurwitz, 1959).

As mentioned in the introduction, a more promising lead comes from Stewart (1959) who interpreted her results in terms of a general theory of reinforcement based on the cortical arousal hypothesis of Hebb (1958) (see also Duffy, 1957; Bindra, 1958). In making predictions about the functional relation between the *intensity* of a stimulus and its *reinforcing* power, Stewart extrapolated from a family of curves relating strength of (escape) responding maintained under fixed-interval reinforcement to noxious stimulus intensity—aversive light—with duration of the fixed interval as a parameter (Kaplan, 1957). In general, the data supported Kaplan's (1957) set of curves, although Stewart suggests that the maxima for the lower intensities exceed the minima of the higher intensities. In other words, for low ratios of reinforcements, low intensities will yield a higher response rate than for high intensities of the stimulus.

Further research requires to be done to substantiate the Kaplan-Stewart hypothesis concerning the functional interaction between reinforcer intensities and schedules of reinforcement. An obvious difficulty arises: in a free responding situation successive reinforcements presumably alter the subjective intensity value of the reinforcer, the limiting case being where the reinforcer results in total stimulus satiation (Glanzer, 1959), for example, like providing *all* the food an animal can eat. For this reason the approach of Bevan & Adamson (1960) which takes into account the adaptation to the stimulus source and attempts to quantify the effectiveness of the reinforcer in terms of the relative difference between background stimulation, reinforcer stimulation and an adaptation process to the reinforcer, has much to recommend it. In the present instance, this theory mediated the prediction of the experimental results obtained. We reasoned that the exposure of subjects to an illuminated box during operant sessions would result in the formation of an internal referent which would attenuate the effectiveness of the light-onset reinforcer. In the case of subjects exposed to a dark box during operant sessions, on the other hand, the reinforcer would function optimally and result in a high response rate. One implication of the theory is that with continued light-onset reinforcement sessions in the dark box, subjects operant tested in the illuminated box and subjects operant tested in the dark box should match their bar-pressing response rates, since the dark-tested subjects should be adapting to the light-onset reinforcer and the illuminated-tested subjects should be adjusting to a dark-box with the result that the reinforcer gains in effectiveness. Our evidence is inconclusive: group A reduced their mean response rate from 6.2 responses per min. on the first light-onset session to 3.1 responses per min. on the last session; group B reduced their mean response rate from 2.9 responses per min. to 1.2 responses per min.

REFERENCES

- APPEL, J. B. & HURWITZ, H. M. B. (1959). Studies in light reinforced behaviour. IV. Effects of apparatus familiarisation. *Psychol. Rep.* 5, 355-6.
- BEVAN, W. & ADAMSON, R. Internal referents and the concept of reinforcement. *Proceedings of the Second Air Force Interdisciplinary Behavioral Sciences Conference* (ed. Washburne, N. F.). (In the Press.)
- BEVAN, W. & ADAMSON, R. (1959). Reinforcers and reinforcement: their relation to maze performance. *J. exp. Psychol.* 59, 226-32.

- BINDRA, D. (1959). *Motivation: A Systematic Reinterpretation*. New York: Ronald Press.
- COCKRELL, J. T. (1952). Operant behavior of white rats in relation to the concentration of a non-nutritive sweet substance used as reinforcement. Ph.D. dissertation. Indiana University.
- COLLIER, G. & SISKEL, M. (1959). Performance as a joint function of amount of reinforcement and inter-reinforcement interval. *J. Exp. Psychol.* **57**, 115-20.
- DUFFY, E. (1957). The psychological significance of the concept of 'arousal' and 'activation'. *Psychol. Rev.* **64**, 265-75.
- GIRDNER, J. B. (1953). An experimental analysis of the behavioral effects of a perceptual consequence unrelated to organic states. Ph.D. Dissertation, Duke University.
- GLANZER, M. (1958). Curiosity, exploratory drive and stimulus satiation. *Psychol. Bull.* **55**, 302-315.
- GUTTMAN, N. N. (1953). Operant conditioning, extinction and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *J. Exp. Psychol.* **46**, 213-24.
- HEBB, D. O. (1958). *A Textbook of Psychology*. Philadelphia: Saunders.
- HURWITZ, H. M. B. (1956). Conditioned responses in rats reinforced by light. *Brit. J. Anim. Behav.* **4**, 31-3.
- HURWITZ, H. M. B. & APPEL, J. B. (1959). Light onset reinforcement as a function of the light-dark maintenance schedule for hooded rats. *J. Comp. Physiol. Psychol.* **52**, 710-12.
- KAPLAN, M. (1957). Intensity escape functions under interval schedules of reinforcement. *Psychol. Rep.* **3**, 251-4.
- MARX, M. H., HENDERSON, R. L. & ROBERTS, C. L. (1955). Positive reinforcement of the bar pressing response by light stimulus following dark operant pretests with no after effect. *J. Comp. Physiol. Psychol.* **48**, 73-6.
- ROBERTS, C. L., MARX, M. H. & COLLIER, G. (1958). Light onset and light offset as reinforcers for the albino rat. *J. Comp. Physiol. Psychol.* **51**, 575-9.
- ROBINSON, J. S. (1959). Light onset and termination as reinforcers for bar pressing in rats living under normal laboratory light conditions. *Psychol. Rep.* **5**, 793-6.
- STEWART, J. (1959). Effects of light intensity and reinforcement scheduling of the positive and negative properties of light stimulation. Ph.D. Dissertation, University of London.
- STEWART, J. (1960). Reinforcing effects of light onset as a function of intensity and reinforcement schedule. *J. Comp. Physiol. Psychol.* **53**, 187-93.

(Manuscript received 30 November 1959)

A PRELIMINARY NOTE ON THE BROOK REACTION TEST

By D. F. BROOK* AND A. W. HEIM

From the Psychological Laboratory, University of Cambridge

An experiment is described on the assessment of interests. The underlying hypothesis is that people's individual interests will be revealed by their responses to an ambiguous word-association test. The stimulus words chosen had two or more meanings, of which at least one related to a specific interest. Each of the thirty subjects was interviewed immediately after taking the test and was rated on nine predetermined interests, using a 4-point scale. Significant correlations between the two criteria were obtained for seven of the nine interests.

I. INTRODUCTION

Several attempts have been made by psychologists to find means of assessing the interests of adults and children. Interests play a major role in education at all levels, in vocational guidance and selection, in problems of social psychology and, to some extent, in the clinical field. The best known work is probably that of Strong (1927, 1933) whose *Vocational Interest Blank* has been widely used. The Allport-Vernon Study of Values (1931) and Kuder's *Occupational Interest Inventory and Preference Record* (1934) come also to mind. In all of these, the subject is required to answer a questionnaire in which the relevant interests are discernible fairly clearly by the subject and the technique involves multiple, and usually 'forced', choice. The subject has to choose between x , y and z —even if he feels, perhaps, that he likes (or dislikes) them all equally.

Such tests have proved over the years to be of some value. There are certain subjects, however, for whom they are unsuitable: for instance, the insightful and unscrupulous candidate in a selection situation, who wants the job and divines which are the more expedient choices to make; the serious-minded obsessional, who finds the 'forced choice' method distasteful or even disturbing; and the sophisticated subject who mistrusts the directness of the approach and complains that he would have responded differently in the past or in other circumstances.

Other attempts have dealt either with a very limited number of interests, as, for instance, Peel's work (Peel, 1956, pp. 159-162) on the practically oriented versus the non-practically oriented child, in which he asked his group of subjects to write as many meanings as possible in reply to certain ambiguous stimulus-words. Or, on the contrary, they are concerned with broad categories of a dispositional, rather than an interest, nature. In this group would be classed certain projective tests and some association word tests, used especially as an aid in differentiating between the normal and the unstable (Jung, 1918).

In this review, the interview must not be forgotten. It is probably the most frequent method of assessing interests, employed both by the psychologist and the layman. Used with care, the interview can be a valuable tool, but its unreliability has been repeatedly demonstrated. Furthermore, too many people turn out on

* Mr Brook died in an air crash over Spain in 1959, shortly after taking his degree. This research project, initiated and supervised by Dr Heim, was carried out by Mr Brook.

interview to be indiscriminately fascinated by every topic under (and over) the sun.

It is felt, therefore, that here is a field which would repay research: that if a technique could be developed combining the virtues of these methods without their drawbacks, it might well prove useful. The purpose of the Brook Reaction Test is to eliminate 'forced choice' and the possibility of line-shooting whilst retaining a high degree of specificity and complete objectivity in administration and scoring.

II. TEST AND AIMS

The test is of the Free Association type. A list of words is read to the subject and he is asked to respond to each word as quickly as possible, with the first word or phrase that enters his head. The stimulus words, however, all have two or more meanings; and the meanings are associated with a variety of different interests. For example, the word 'Train' is of interest to the engineer, the dressmaker, the athlete and the learning-theorist (assuming they are interested in their work) because of four different senses in which the word is used. The hypothesis is that the first replies with some such word as 'engine'; the second with, for instance, 'brocade'; the third with, say, 'rowing'; and the fourth with 'negative transfer'. Subjects offering 'carriage' would have to be questioned further, since they might have in mind either meaning (i) or (ii).

The stimulus words are simple everyday words, familiar to every school child. Since they have in fact two or more quite distinct meanings, it will not in general be the tester's task to *interpret* the results, once the test has been standardized. It should usually be clear which meaning the subject has unthinkingly understood. When any doubt obtains, further (non-leading) questions must follow the testing: and, if the doubt persists, the particular response must be omitted from the score. The test is presented as one of reaction time, with stop-watch in evidence.

The aims of the preliminary experiment were as follows:

(1) To ascertain whether each subject tends to give a fairly consistent pattern of responses. Thus subject P, for instance, would respond to the *athletic* significance of most of the words which offer this possibility among others; subject Q, on the other hand, would usually take the *social* significance of the words.

(2) To determine which words and which interests 'hold up' on the preliminary testing. It was realized that the number and type of particular interests chosen was arbitrary and that a series of trials, alterations and retrials would probably be necessary.

(3) To develop an objective system of scoring.

(4) To validate the test against some external criterion.

III. THE EXPERIMENT

A. Stimulus words and interest categories

At the outset, eleven interest categories were chosen: *academic, aesthetic, outdoor, sport, social, political, practical, religious, scientific, sex* and *entertainment*. A much larger number would have required too long a list of stimulus words, whilst a much shorter number would have covered too narrow a field. Two of the above categories.

however, soon emerged as impracticable: *entertainment* was dropped because it proved to overlap with *aesthetic*, and *sex* was dropped because insufficient validation data were forthcoming. Nine interest categories were left then, which are described and illustrated.

Academic denotes the subject's studiousness, his liking for serious reading and matters primarily theoretical. For example, stimulus word 'Set': subjects giving 'book' in reply are categorized as *academic*, as opposed to those giving, for instance 'group' (*social*) or 'tennis' (*sport*).

Aesthetic covers interest in music, the plastic arts, literature, the theatre. For example, stimulus word 'Canvas': the response 'picture', for instance, is an *aesthetic* response as opposed to 'tent' (*outdoor*) or 'vote' (*political*). Homonyms with different spellings are permissible since the test is given orally.

Outdoor signifies enjoyment of outdoor life, interest in agriculture and the countryside generally. It proved relatively easy to distinguish *outdoor* from *sport*, which denotes interest in active games. For example, stimulus word 'Pass': response 'mountain' is classed as *outdoor*; 'rugger', as *sport*; 'examination', as *academic*. Or stimulus word 'Bail': response 'hay' is *outdoor*, 'wicket' is *sport*.

Social includes conventional social interests as well as psychological and humanitarian ones. For example, stimulus word 'Pique': response 'offended' is classed as *social*; 'Switzerland', as *outdoor*.

The connotation of *political* is rather narrow, in order that it should not overlap with other categories. The illustration 'Canvas(s)' was given above. Another typical stimulus word here is 'Party', its other category being *social*.

By *practical* is intended interest in working with the hands, metal-work, carpentry, model-making, etc. For example, stimulus word 'Vice': response 'woodwork', as opposed to 'virtue' or 'president' (both *social*).

Religious is self-explanatory. For example, stimulus word 'Primate': response 'bishop' is classed as *religious*, 'ape' as *scientific*.

In practice it was usually easy enough to distinguish between *scientific*, *practical* and *academic* responses. Again, an example may make this clear. Thus with the stimulus word 'Angle', the response 'geometry' was classed as *scientific*; 'joint' as *practical*; and 'Saxon' as *academic*.

At an early stage, the list consisted of 106 stimulus words, each of which connoted at least one meaning relevant to one interest category. The order of these words was randomized. In order to allow for the possible tendency for one word to determine to some extent the response to subsequent words, the list was presented to alternate subjects in the order 1-106 and 106-1.

B. Subjects and testing procedure

In this preliminary investigation, thirty Cambridge students took part, of whom twenty-four were men and six were women. There were, perhaps unfortunately, twice as many Arts as Science students.

Subjects were seen individually. They knew nothing about the object of the inquiry. They were told that the experimenter would read aloud a list of words; that they should respond with the first word or phrase that came naturally to mind; and that their speed of response was of major importance. This served two purposes.

- (i) It encouraged the subject to hurry and think about hurrying, so that he was less likely to be critical of his own responses or to ponder over the stimulus words.
- (ii) It convincingly disguised the purpose of the test.

At the end of the test, the experimenter went over any doubtful responses, for instance, those whose relation to the stimulus word was unintelligible or ambiguous (as in 'carriage' above). He also welcomed any general comments volunteered by the subject. A semi-formal interview followed, the total session usually lasting about 40 min.

C. *The interview*

This partially structured interview was given in an attempt to validate the test. The interview, in which all subjects participated willingly, but without being too deeply ego-involved, was based on a list of questions which was short and factual, and did not intrude unduly on the subject's sense of privacy. The proforma embodying these questions is reproduced in the Appendix.

Subjects naturally varied in their degree of expansiveness. All proved reasonably co-operative, however, and the experimenter had no cause at any time to doubt their honesty or good faith.

The data obtained from each subject about each interest category (I.C.) was expressed in terms of a 4-point rating scale: 0, 1, 2 and 3. Thus, a subject revealing no interest in sport, for instance, would be rated as 0; one showing little, would be rated as 1; a subject with more sporting interests would gain 2; the real enthusiast would rate 3; and so on with each of the I.C.'s. Each subject was given an I.C. rating on each of the nine categories.

IV. RESULTS

A. *Treatment*

Inspection of the association word test data revealed a number of responses which had no obvious bearing on any of the predetermined I.C.'s. These included irrelevant and highly obscure answers and some clang responses. Fruitful though these might prove for a different type of test, it was decided to eliminate stimulus words producing many such responses. The words intended to relate to *entertainment* and *sex* were also removed. This left a total of sixty-five useful words.

The experimenter examined each of these in turn and counted the number of different interest responses accorded to each word. The word 'Arc', for instance, prompted nineteen *religious* responses, eight (mathematical or) *scientific* responses and a few miscellaneous responses such as 'royal' or 'de Triomphe'. In the case of stimulus word 'Eden', twenty-two people responded in a *religious* manner ('garden', 'apple', etc.) and four, *politically* ('Suez', 'Macmillan', etc.).

Since thirty subjects participated, the number of times a word could elicit a particular response ranged from 1 to 30. This range was divided into six groups: 1-5, 6-10, 11-15, 16-20, 21-25, 26-30. It was argued that the more infrequent the meaning taken by members of the group, the keener the interest for the individual. In the example of 'Eden' given above, for instance, the four giving *political* replies are likely to be *more* interested in matters political than the twenty-two *religious* repliers are in matters religious. For this reason, the above groups were, experimentally, scored as 6, 5, 4, 3, 2 and 1, respectively. Thus when twenty-six or more subjects

gave a particular interest response to a word, this was scored as 1; when twenty-two (as in *Eden-religious*, above), this was scored as 2; when four (as in *Eden-political* above), this was scored as 6.

It was felt that the exceptional subject who gave a 'miscellaneous' response to a stimulus word which elicited the *same* specific interest response from over two-thirds of the group might well be positively antipathetic to that interest. The decision was made, therefore, to deduct one point from the relevant interest score of such a subject, when 21-25 of the others were united on one interest meaning, and to deduct two points when the majority was over 25.

By means of these two principles, it was possible to compute a score for each subject on each interest category.

B. Findings

Comparison of test results with interview results is expressed in two ways: first, in terms of correlation coefficients, and secondly, in graphical form. Table 1 shows the results of correlating the test score with the interview ratings as made on the 4-point scale, using Kendall's *tau*.

Table 1. *Correlations between word association scores and interview ratings*

Interest category	<i>tau</i>	<i>P</i>
Sport	0.63	0.01
Religious	0.50	0.01
Scientific	0.49	0.01
Political	0.44	0.01
Practical	0.40	0.01
Outdoor	0.38	0.01
Aesthetic	0.35	0.05
Academic	0.15	Not sig.

It will be seen that eight out of the nine I.C.'s have been included. The ninth, the *social* category, does not appear because it was found that point 0 of the interview rating scale had never been used for this interest: everyone was rated 1, 2 or 3. Since this meant that in effect a 3-point scale only had been employed, this category is omitted here.

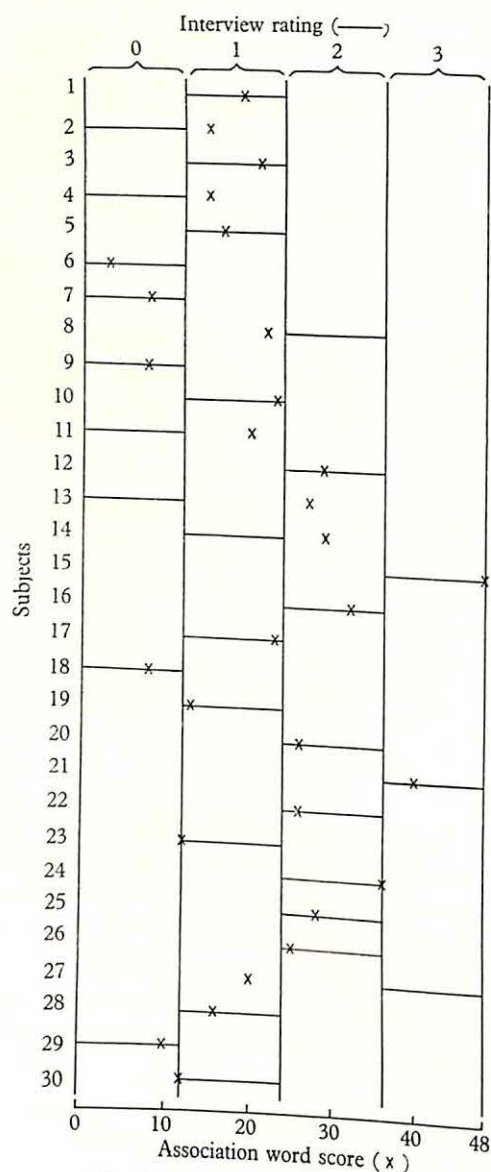
It may be seen that the *tau* correlations between interview data and test score are significant in seven out of the eight cases, reaching the 0.01 level of confidence for six interests.

The second method of presentation is more complicated but it yields considerably more qualitative information. For reasons of space, graphs are shown for four interests only: *sport* and *scientific*, both of which show a close relationship between the Word Reaction Test and interview assessment; *outdoor*, which shows a weaker relationship; and *aesthetic* which is the weakest of those given.

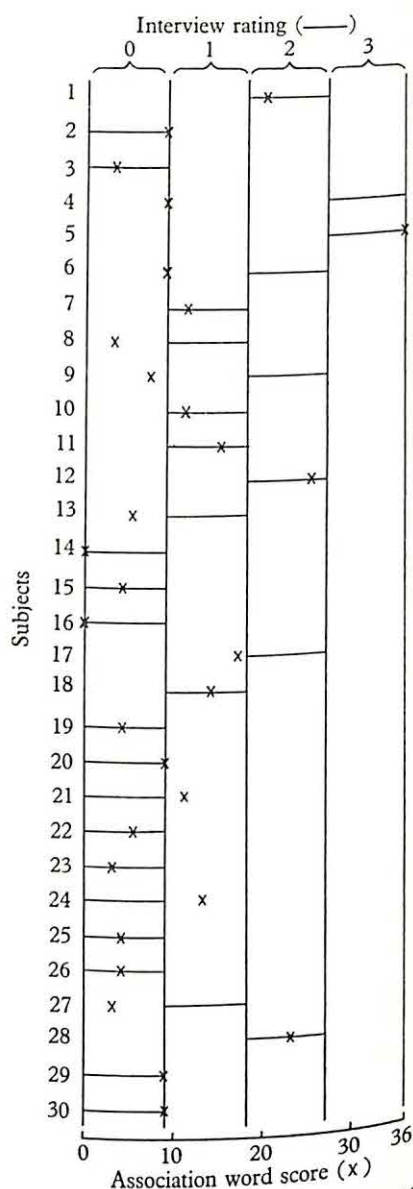
The numbers on the left represent the thirty subjects in the order in which they were tested. The lower scale refers to the association word scores: the higher the score the more intense the interest. The crosses show the association word score gained by each subject. Thus, for instance, on *sport* subject 29's score was 10, while subject 30's score was 12. It may be seen that the highest *sport* word-score was obtained by subject 15 who gained 48 points.

The interview ratings, 0, 1, 2 or 3, are expressed by the horizontal lines, as shown

on the top abscissa. Their length varies from I.C. to I.C. because the range of word score, from 0 to the highest score gained, was always divided into four equal parts. and the highest score was naturally not identical for all the interests.



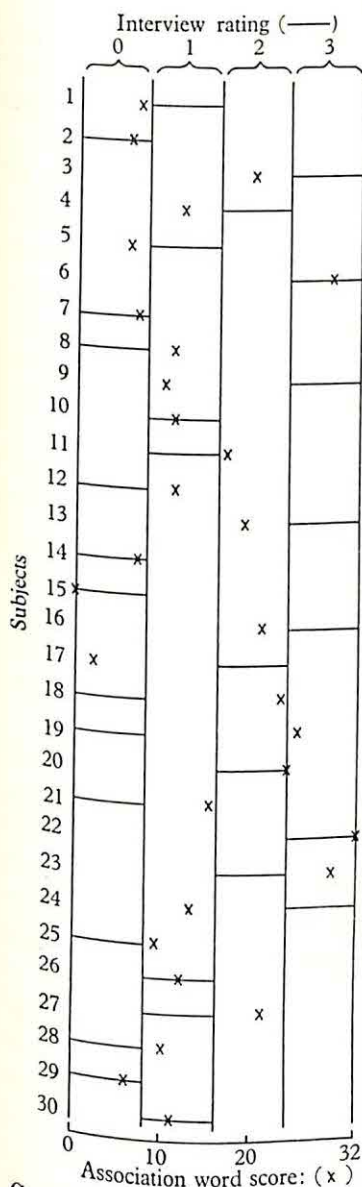
Graph 1. Interest category: *sport*.



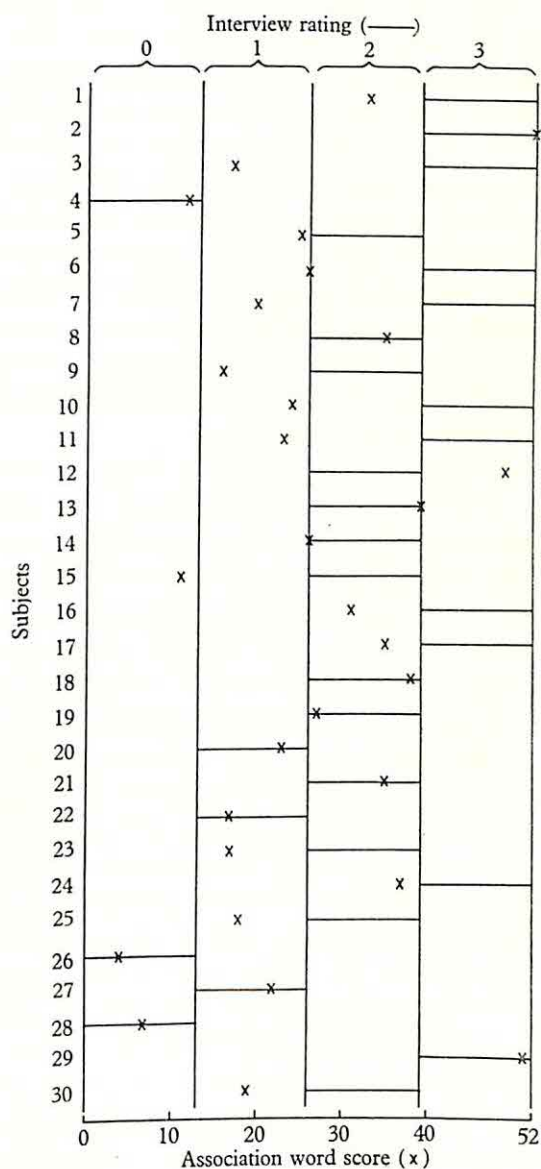
Graph 2. Interest category: *scientific*.

Thus, a cross falling on a line, indicates close agreement between test score and interview rating; when the cross and line just fail to meet, agreement is less; and when the cross and line are a long way apart (as for subject 27 on *sport*) agreement is lacking. As a rough estimate, we may count the number of times, out of the possible 30, the line and cross made contact. On *sport*, this occurred 23 times; on *scientific*, 21 times; on *aesthetic*, 14 times; and on *outdoor*, only 11 times. Moreover, the discrepancies between cross and line are larger, as well as more frequent, on the latter two I.C.'s: see, for instance, subjects 17, 18 and 19 on *outdoor*, all of whom behaved

'appropriately' on *sport* and *scientific*. At present, the question remains open as to whether in such cases the association words are ill-chosen, the interest is poorly defined, or the interview less diagnostic regarding some interests than others.



Graph 3. Interest category: *outdoor*.



Graph 4. Interest category: *aesthetic*.

V. DISCUSSION

The test is clearly still in its infancy but at this preliminary stage, it looks promising. The relations between test and interview and between the reactions and introspections of the subjects, proved very much what had been hoped. Subjects were evidently unaware of the purpose of the test (and interview); they accepted the situation as testing speed of reaction; and none of them realized that all the words were ambiguous—a few subjects actually stopping the experimenter in the course of the test to point out that 'the last word has two meanings'.

The first six categories in Table 1 appear satisfactory, though their diagnostic value could probably be improved by increasing the number of relevant words and perhaps rejecting some of the less discriminatory words.

There are some data which throw light on the relative failure of the last three I.C.'s, *aesthetic*, *academic* and *social*. For example, the experimenter tended in the interview consistently to overestimate the extent of aesthetic interest. (Inspection of Graph 4 reveals that in almost all the discrepancies, the interview line is to the right of, i.e. 'higher than', the test cross.) This was due largely to the width of the category. If a subject showed great interest in one specific aesthetic experience only, such as painting or the drama, he was rated high on the interview scale. In the test, however, he would be liable to gain a low score since his particular interest would be relevant only to a small proportion of the *aesthetic* words. The *academic* category was represented by insufficient stimulus words. It also overlapped the *aesthetic*, especially with regard to literature. It might be a good idea to introduce a separate interest category for literature. This should clarify both the *academic* and the *aesthetic* categories considerably.

Social was unsatisfactory, largely owing to confusion between interest in people as individuals on the one hand, and interest in ideas about people or simply the possession of a broad humanitarian outlook, on the other. The latter was found in fact to be closely related to religious interest, but this was not necessarily true of the former. Subjects intimating humanitarian interests on interview were (sometimes mistakenly) rated highly on *social*. Again, people who joined a political organization tended to be rated high on *political*, in the interview. But subsequent discussion revealed that such joining often reflected primarily *social* interests.

Thus, discussion with subjects after the end of the experiment threw light on several anomalies and suggested that where discrepancies obtained, the association word test tended to give more accurate results than the interview rating. In view of the objective fact of ambiguity in the stimulus words used, the Brook Reaction test is a far less subjective means of assessment than is the interview. An example may illustrate this point. Suppose the word 'Wall' had been used in the test: this might evoke (i) 'fives' (*sport*), (ii) 'hedge' (*outdoor*), (iii) 'painting' (*practical* or *aesthetic*), or (iv) 'flower' (*outdoor* or *social*). We do not use such a word since these four hypothetical subjects would all be responding to the same meaning—the only meaning possessed by the word 'Wall'. There is an element of near-conscious choice which is present in all free associations, whether the stimulus words are ambiguous or unambiguous. The I.C. of 'painting' and 'flower' could be ascertained only by having recourse to introspection. As indicated above, this is rarely necessary with the Brook Reaction Test, since the meaning automatically taken by the subject is usually self-evident. One of the subjects in reply to 'Motor' gave the response 'senses': his implicit understanding of the word was different from that of the majority who tended to give 'car', 'machine', etc. The basic decision as to category is made—and made unwittingly—by the subject and not the tester.

The test is not concerned with neurosis or even neuroticism. Long reaction times were noted but they were rare, for care had been taken to avoid words which have obvious emotional or embarrassing connotations. There were a few clang responses but no silences. Most of the subjects produced swift, direct 'verbal' responses and

found the test easy and pleasant. A small minority, however, complained of difficulty and were less spontaneous in their responses. These latter subjects described visual, rather than conceptual, reactions to the stimulus words. One described graphically how the word 'Party' conjured up a vivid scene dominated by balloons, paper hats and bottles, from which a word had to be selected. No noticeable difference was found, however, in the results of these two types of subject. Nor did the variation in order (1-106 and 106-1) yield any significant differences.

In addition to increasing the number of categories and stimulus words, and purifying these, it is proposed to try the test on other groups including school children, graduates and non-university adults. It is hoped also to obtain a further group of students more evenly balanced in regard to faculty and sex. It would be an advantage if in future work, test and interview could be conducted by two different experimenters, or, perhaps, two experimenters could sit in at one interview and make independent ratings. The interview might be supplemented by a self-rating scale on the relevant interests. Finally, the possibility of other types of validating criteria should be explored.

APPENDIX

Proforma used in interview

Name

School or University

Age

Date

1. What newspapers, if any, do you
 - (a) take at home,
 - (b) read at home,
 - (c) read at university?
2. What magazines or periodicals, if any, do you
 - (a) take at home,
 - (b) read at home,
 - (c) read at university?
3. How did you spend your last two weekends?
4. What did you do during your last two vacations?
5. How do you get hold of most of the books that you read?
6. If you belong to a public (non-faculty) library, which sections do you most use?
7. What are you reading at University?
8. What were
 - (a) your favourite subjects at school?
 - (b) those you most disliked?
9. Have you been in the last two months to the theatre, cinema, opera, ballet or concert, museum or art gallery?
10. Do you belong to any clubs or societies?
11. Do you find that you have the time and the facilities for sport?
12. What are your favourite wireless or T.V. programmes?
13. Do you have any particular outside interest?
14. What do you hope to do when you leave college?

REFERENCES

- ALLPORT, G. W. & VERNON, P. E. (1931). A test for personal values. *J. Abnorm. Soc. Psychol.* **26**, 231-48.
- JUNG, C. G. (1918). *Studies in Word Association*. London: Heinemann.
- KUDER, G. F. (1934). *Preference Record—Vocational*. Chicago: Science Research Associates.
- PEEL, E. A. (1956). *Psychological Basis of Education*. Edinburgh: Oliver and Boyd.
- STRONG, E. K. JR. (1927). *Vocational Interest Blank for Men*. Palo Alto, Calif: Consulting Psychologists' Press.
- STRONG, E. K. JR. (1933). *Vocational Interest Blank for Women*. Palo Alto, Calif: Consulting Psychologists' Press.

(*Manuscript received 7 December 1959*)

A RESULT IN VISUAL AESTHETICS

By CHRISTOPHER ALEXANDER

*(Joint Center for Urban Studies of the Massachusetts Institute of Technology
and Harvard University)*

Subjects were given eight forms and asked to sort them in a number of ways on the basis of overall similarity; they were also asked to state the order of their preferences among the forms. From the data generated, four tentative results emerged:

(1) By giving ourselves an appropriate verbal set, we can make ourselves see (and categorize) a group of forms in many different ways. It appears, however, that there is a natural way of categorizing them which is independent of any verbal set, and which depends on the formation of visual non-verbal concepts.

(2) Preference for the forms is dependent on the way the forms are seen—on the visual similarity dimensions. But the dependence is incomplete. It might perhaps be better called a linkage—comparable to the linkage between hue and brightness.

(3) The peculiarly weak nature of this linkage suggests the hypothesis that the beauty of a form cannot be explained in terms of any visible qualities or attributes that it has, but only in terms of the operations performed in the brain of the observer.

(4) Aesthetic discrimination is independent of all other kinds of perceptual discrimination.

I. INTRODUCTION

The way in which we look at things lies at the heart of visual aesthetics.

It is possible to look at a form in many different ways. We look at different aspects of the form—or we look for different things in the form. We may notice its plasticity, its movement, its simplicity, or any of a hundred other characteristics—characteristics that we can concentrate on one at a time.

To look at a form in a certain 'way' is to pay attention to (or to look at) a particular characteristic of the form. Now, we can set ourselves to look at any characteristic we wish. But what if we do not set ourselves deliberately at all? How do we look then? This is the problem we are interested in. We wish to find out:

- (1) the 'ways' in which a subject naturally looks at forms;
- (2) whether there is a connexion between his liking for the forms and the ways in which he looks at them;
- (3) if there is such a connexion, its nature.

For the vague notion of a 'way' we shall substitute one more suitable for operational definition and analysis: the two-ended dimension. The situations where a subject pays attention to the plasticity, the movement, the simplicity of a form, we shall describe by saying that he is using the dimensions 'plastic-flat', 'dynamic-static', 'simple-complex'.

The first problem, then, is to find out which dimensions best describe the way a subject looks at forms. We could try several techniques.

(1) We might ask him to say which characteristics of the forms he paid attention to. And this would involve his stating, *in words*, the dimensions he believed himself to use. We could call them his introspective dimensions.

(2) We might present him with a long list of dimensions (plastic-flat, dynamic-static, simple-complex, open-closed, and many others), and ask him to place a number of forms on each of them (according to their plasticity, their movement, their sim-

plicity, their degree of closure, and so on). The semantic space set up by these dimensions could be factored and redefined in terms of a minimal set of dimensions (Osgood *et al.*, 1957; Tucker, 1955), which might then be said to describe the way the subject looked at the forms.

(3) More subtly still, we might use the following procedure, known as the method of triads. Three forms are shown to the subject, and he is asked which two are most alike. He is then asked to say in what respects the two are alike, and how the third one differs from them (Kelly, 1955). After he has given such answers for a number of triads we can see which aspects of the forms he pays most attention to when making his judgements, so we can construct (using, if possible, words that the subject himself has used) a set of dimensions that describe the way he looks at things (Henderson, Kates & Rohwer, 1959).

Yet all these methods are verbal ones—attempts made to fit verbal categories to visual phenomena. And while this is by no means impossible (it has been done with some success by critics, after all), there is a good deal of evidence to show that such attempts cannot get to the 'heart' of visual aesthetics.

Before we see why this is so, in full, let us consider a single incident that occurred during a triad experiment. A 12-year old girl was shown a Canaletto view of St Mark's Square, a Guardi view of the Grand Canal, and a line drawing of a single boat by Corot (all on postcards). Immediately she put the Venetians together as the two that were most alike—they were in fact extremely similar. But suddenly she remembered that she had to explain why or in what respects they were alike. And promptly she changed her mind, put the Guardi and the Corot together, and said, 'Those two, they've both got boats on them'. Visually this pairing was ridiculous. She had been forced by the demands of the experiment to group them in a way she had a word for. The distinguishing quality of the Venetian paintings was too hard for her to explain, though she could see it very well.

It is held by some psychologists that all our seeing is based on verbal categories (Brown, 1956; Whorf, 1941). That we only see what we have words for, that there is no visual concept formation, only verbal. And such a psychologist would say of the child, 'She actually saw like that; she saw in terms of the words she knew'.

We shall be able to show later that this view is false. That we can see independently of the words we know. For the present we shall simply remember that the child's *first* instinct was to put the Venetian paintings together.

Every experiment in visual aesthetics that deals with words is handicapped. People will not respond according to what they see, but according to the hopelessly inadequate vocabulary they have. And the results will be, just as the above one was, quite valueless from the point of view of visual aesthetics. If we are to achieve interesting results, we must let the subjects use their eyes. We must collect data which reflect only visual behaviour: we must find out the ways in which someone sees without letting him consider even, the words that described his ways of seeing.

II. THE EXPERIMENT

(1) *Stimuli*

The stimuli were 3 × 5 in. white filing cards, each with a single form drawn on it in black ink. About fifty forms were drawn quickly and freely; eight were then chosen from them. They were

chosen for their mutual similarity, for their obviously set-like character, for the fact that they were variants on one another 'in several different directions'. The forms were named, at random, A, B, C, D, E, F, G, H (see Fig. 1). At no point of the experiment was the subject told the names of any of the forms—in case the names led him to make decisions on grounds connected with the letters. Subjects sometimes asked to see the forms upside down and from different angles. This was forbidden. If the subjects had seen the forms from several angles they might have noticed new aspects that would have led them to look at the forms in changed ways when it came to the experiment.

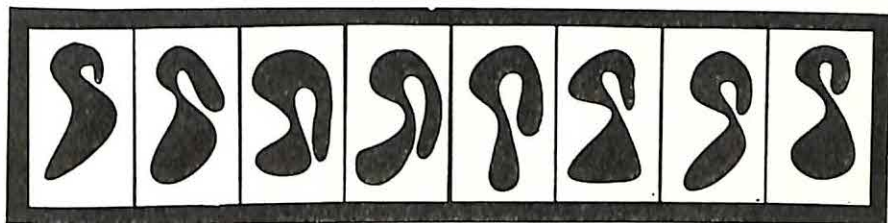


Fig. 1. The stimuli A, B, C, D, E, F, G, H.

(2) Subjects

There were six subjects, all well educated, all used to looking at things for pleasure, all between 20 and 30, three men and three women.

(3) Procedure

Two of the eight cards were chosen at random by the experimenter. They were laid in front of the subject, on a neutral ground. (The decision as to which card should go to the right and which to the left, was also made at random, and noted. During the repeat of the experiment, 24 hr. later, the pair was presented the opposite way round.) The remaining six cards were given to the subject, with the following instructions:

'Look at the six cards in your hand one by one, comparing them with the two cards on the table. For each one you must decide whether it looks more like the right hand or the left hand of the pair on the table. When you have decided, lay it on the table too, to the right or the left, according to the similarity you have observed.' (As the cards were laid down, they were kept at some distance from the pair already there; so that there should be no effects of visual proximity which might prejudice other decisions.) 'When you have placed all six cards in this way, make a check. If you wish to change your mind about any of them you may do so. You are under no compulsion to split the cards evenly; sometimes you may even want to put all six on the same side. Let it depend only on your feelings about the similarity of the forms.'

It was repeatedly made clear to the subject that he should try to be visually naïve, and should avoid making decisions on intellectual grounds. 'In the more difficult cases you will find yourself thinking hard about the decision. You may even feel that you could put a card on one side according to one criterion and on the other side according to some other criterion. If this does happen, stop thinking about it at once. Look away. When you look back, remember you are doing it on overall similarity, because forms look or feel alike. Forget that you are involved in an experiment, and imagine that an acquaintance has suddenly asked you, quite informally, "Which one does this look more alike?" Decide quickly; just like that.'

The visual interaction of some of the forms, when they were seen close together, seemed to have a disturbing effect also. To avoid it, where X and Y were forms on the table, and O one of the six forms being matched against them, subjects were given this instruction: 'O is to be compared with only one of X and Y at a time. Thus, while you look at O and X, keep Y covered. And when you look at O and Y, cover X. You are thus always looking at a pair, (O, X) or (O, Y). And you are to decide which

makes the "closer" pair, (O, X) or (O, Y). What was in fact being investigated here was the relation between a number of perceptual distances. If the subject found O more like X than Y, we may express this by saying that, for him, O is nearer to X than it is to Y; the perceptual distance OX is smaller than the distance OY.

As a further precaution the subject was asked never to put two cards really close together, but to keep several inches between them. When cards are very close to one another, the shape of the white between the two black forms becomes very important to the eye—and may upset the judgements made.

After the subject had made his six decisions they were written down like this:

X	Y
O	O
O	O
	O
	O

The subject went through this sorting procedure for each possible pair, i.e. twenty-eight times. The session took between an hour and an hour and a half. Twenty-four hours later all twenty-eight were repeated, in a different order, and with the pairs laid on the table which ever way round they had not been on the first occasion.

Subjects were quite often not consistent. For those cards which were placed in the pile they had been in 24 hr. before, the results were accepted, since here the subject seemed fairly certain, and there was very little doubt about his feelings on the matter. For those cards which were put in the pile they had not been in the first time, however, a closer scrutiny was necessary. Such inconsistency might have been caused in one of three ways.

(1) The most important cause, undoubtedly, lay in the very way the similarity judgements were made. It may be more likely that a subject puts O with X than that he puts it with Y. But if the decision is a hard one, as it often was, we shall at best observe this greater likelihood as a greater frequency. He will put O with X more often than with Y. And this can lead us to say that OX is smaller than OY by a narrow margin. (If the frequency turned out to be 50-50, we should have to call these two perceptual distances equal.) So if the subject put a card first in one pile, and then in the other, he was asked to sort it yet again—at once; and then again. Until it was clear which pile better pictured his feelings on the matter. (Each time he resorted to the cards were rearranged so that he should not remember what he had done before.)

(2) The second possible reason for a subject's inconsistency lies in the arrangement on the table. The subject may be inclined to put cards to one side of his body rather than the other. Normally this factor will cancel out in the end, under the constant rearrangement. But if there is a card which he puts always to the same side, regardless of which pile lies on that side, one must assume him to be indifferent to the choice between piles and governed only by his left-right preference; and the perceptual distances will have to be taken as equal. (In all the experiments conducted, this happened only twice.)

(3) Thirdly, he may appear inconsistent because during the first time through he was unfamiliar with the cards, and has changed his mind now that he knows them better. This seemed to be the case a good deal of the time; during the second session

the subject was much surer of his attitude towards the cards than during the first. Where this was so, further sorting tended to support the subject's second decision rather than his first.

It is interesting to note that while subjects disagreed widely in their sorting on the first time through, there was strong agreement about the final versions. This suggests that the subject settled down gradually to his balanced judgement; and, what is more, it suggests that this balanced judgement corresponds only to stimulus characteristics, and independent of the particular subject whose judgement it is. Perhaps, if one were to stretch the experiment out even further, and let subjects settle down to their opinions still more slowly, we should find still greater agreement.

During the second session the subject was given a further task. Each time he finished with a pair, he was asked which member of the pair he preferred. Every subject thus made twenty-eight paired comparisons—which generated an order of preference over the eight cards. It was done during the second session only, so that the subject should be thoroughly familiar with the forms by the time he came to state his preferences. (In every case but one these paired comparisons were quite consistent and led to no intransitivities. The one subject (no. 7) who did produce intransitive results also produced very odd similarity data. He found all the tasks difficult, and said he could only do them on a consciously intellectual basis. His data have therefore not been included in the results of the experiment.)

Subjects were asked to make their preference judgements on grounds of shape alone—not to pay attention to associative overtones, but to judge the shape purely as a shape. Of course, they were not able to do this. The considerable disagreement illustrates this quite clearly. One always reads forms in a certain way. But they were trying to do it *this* way; and their judgements were governed by the feeling for form as much as possible.

Finally, when the whole experiment was over, the concept of a dimension was explained to the subject (though no specific cards were mentioned as illustrations, in case they biased his answer). The cards were now laid out in front of him, in random order, and he was asked what he thought he had been using as the bases for his similarity decisions, in spite of the fact that, at the time, he had been asked expressly to use no bases, principles, or criteria.

Each basis he gave—like longnose—shortnose—he was asked to illustrate with the most extreme examples. Thus he would be asked to point to the form with the longest nose and to that with the shortest. He was giving here his introspective dimensions; and most subjects gave three or four.

III. TABULATION AND ANALYSIS OF DATA

For each subject we now have the following data:

(1) Twenty-eight tables of the form

C	E
D	A
F	B
	G
	H

(2) A preference order generated by paired comparisons.

(3) A list of introspective dimensions given by the subject to account for his sorting behaviour.

Each table like

C	E
D	A
F	B
	G
	H

is in fact a condensed statement of six inequalities among the perceptual distances; for since F is put under C rather than under E, the table indicates that $CF < EF$; and five other facts of the same kind. Similarly the tables

C	F	and	E	F
B	A		C	A
D	G		D	B
E	H		G	H

tell us that $CE < EF$ and $CE < CF$. We may combine the three inequalities to give $CE < CF < EF$. (When it happened—as it did about once for every subject—that the three inequalities were inconsistent, then one of them was reversed; whichever one the subject had been most uncertain of, whichever one he had changed his mind about most often.) The statement $CE < CF < EF$ tells us that in the triad CEF the perceptual distance EF is the greatest of the three, so we write the triad ECF. The position of C between E and F is most important and will be referred to as this triad's betweenness.

Our data give us betweenness information of this kind for every one of the fifty-six triads; and for every subject (see Fig. 2). As has been shown in a recent paper (Hays, 1959; summarized in Coombs, 1958), this information as to betweenness may be used to generate orders which are closely similar to the dimensions obtained in factor analysis. (Factor analysis itself is not possible since there are no cardinal data available.) The following is a brief outline of the procedure used to generate these 'dimensions'. It is to be found, in full, together with its mathematical justification, in the papers mentioned. It is of course carried out separately for each subject's data.

(1) Select, by inspection of the data, that pair which seems to be most dissimilar (where the perceptual distance is the greatest), and use this pair as end-points of the first dimension.

(2) Order the remaining six letters between these end-points so as to accommodate as many of the fifty-six triads as possible. That is to say, construct that order which preserves betweenness for as many of the triads as possible.

(3) Repeat the above procedure for those triads not accommodated by the first dimension. If these are not enough to define a second dimension uniquely, construct that one which overlaps the first as little as possible. (It is inevitable, in spite of this, that there will be some triads whose betweenness is satisfied on both dimensions.)

In principle the extraction of dimensions should go on in this way until all the triads have been accommodated. But, in fact, no more than two dimensions were ever needed. All the triads, bar one or two, were accommodated by the first two

Triad	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6
ABC	ABC	ABC	ABC	ABC	ABC	ABC
ABD	ABD	ABD	ABD	ABD	ABD	ABD
ABE	ABE	ABE	ABE	ABE	ABE	ABE
ABF	ABF	ABF	ABF	ABF	ABF	ABF
ABG	ABG	AGB	AGB	AGB	AGB	AGB
ABH	ABH	ABH	ABH	ABH	AHB	ABH
ACD	ADC	ADC	ADC	ADC	ADC	ADC
ACE	AEC	ACE	ACE	AEC	AEC	AEC
ACF	AFC	AFC	AFC	AFC	AFC	AFC
ACG	AGC	AGC	AGC	AGC	AGC	AGC
ACH	AHC	AHC	AHC	AHC	AHC	AHC
ADE	ADE	ADE	ADE	ADE	AED	ADE
ADF	ADF	ADF	AFD	AFD	AFD	AFD
ADG	AGD	AGD	AGD	AGD	AGD	AGD
ADH	ADH	AHD	AHD	AHD	AHD	AHD
AEF	AFE	EAF	AFE	AFE	AFE	AFE
AEG	AGE	AGE	AGE	AGE	AGE	AGE
AEH	AHE	AHE	AHE	AHE	AHE	AHE
AFG	AGF	AGF	AGF	AGF	AFG	AGF
AFH	AHF	AHF	AHF	AFH	AFH	AFH
AGH	AGH	AGH	AHG	AGH	AHG	AGH
BCD	BDC	BCD	BDC	BDC	BCD	BDC
BCE	BEC	BCE	BEC	BEC	BEC	BEC
BCF	BFC	BCF	BFC	BFC	CBF	BFC
BCG	BGC	CBG	CBG	BGC	BGC	CBG
BCH	BHC	BCH	CBH	BHC	CBH	CBH
BDE	BDE	BDE	BDE	BED	BED	BDE
BDF	BDF	DBF	DBF	DBF	DBF	DBF
BDG	BDG	BGD	DBG	BGD	DBG	BDG
BDH	DBH	DBH	DBH	DBH	DBH	BDH
BEF	BFE	EBF	EBF	EBF	EBF	BFE
BEG	BGE	BGE	EBG	EBG	EBG	BGE
BEH	BHE	EBH	EBH	BHE	EBH	BHE
BFG	BGF	FBG	BGF	BGF	FBG	BGF
BFH	BFH	BHF	BHF	BHF	FBH	BFH
BGH	GBH	BHG	BGH	BGH	BHG	BGH
CDE	CED	DCE	CDE	CDE	CDE	DCE
CDF	DCF	DCF	DCF	DCF	DCF	DCF
CDG	CDG	CDG	DCG	CDG	CDG	CDG
CDH	DCH	DCH	CDH	CDH	DCH	DCH
CEF	CEF	ECF	ECF	ECF	ECF	ECF
CEG	CEG	ECG	ECG	CEG	CEG	CEG
CEH	CEH	ECH	ECH	ECH	CEH	CEH
CFG	CFG	FCG	CFG	CFG	CFG	CFG
CFH	FCH	CHF	CFH	CFH	CHF	CFH
CGH	CHG	CHG	CGH	CHG	CGH	CGH
DEF	DEF	EDF	EDF	EDF	DEF	EDF
DEG	EDG	EDG	EDG	EDG	DEG	DGE
DEH	EDH	EDH	EDH	DEH	DEH	DEH
DFG	DGF	DGF	DGF	DGF	DGF	DGF
DFH	DFH	DHF	DHF	DFH	DHF	DHF
DGH	GDH	DGH	DGH	DGH	DGH	DGH
EFG	EFG	EGF	EGF	EGF	EGF	EGF
EFH	EFH	EHF	EFH	EHF	EHF	EHF
EGH	EGH	EGH	EGH	EGH	EGH	EGH
FGH	GPH	FHG	FGH	FGH	FHG	FHG

Fig. 2

dimensions extracted.* The dimensions are tabulated in Fig. 3. In Figs. 4a and 4b we have a visual presentation of the same material. Fig. 4a contains the six subjects' first dimensions, and 4b their second ones.

As we can see, the subject to subject agreement as to these dimensions is very high. The coefficient of concordance (Kendall, 1948) is 0.81 for the first dimension, and 0.65 for the second. Both these are above the 0.1 % level of significance. Since the agreement is so good, the modal dimensions are illustrated, for interest's sake, in Fig. 5. These are the principal 'ways' in which the average subject looked at the eight forms shown to him.

Subject	Dimension 1	Dimension 2	Introspective dimensions (in the order they were given)	Preference order
Subject 1	ABGDHFEC	EDGCBFH*	F-B Balanced—likely to topple F-H Sharp—round F-A Masculine—feminine B-A With weight—without weight F-D Pointing left—pointing right	CADGEBFH
Subject 2	ABFGHDCE	EADGBCHF	D-F Thin, linear—fat, solid H-F Round—angular C-A Long nose—short nose	HGAFEBCD
Subject 3	AHGFBCDE	ABDEGHFC	E-A Open—closed A-F Bird—snake F-E Solid—tottery	BGHFDCEA
Subject 4	ABGHFEDC	EDBGACHF	A-H Angular—round H-F Vertical—slanted G-A Indented—not indented	GCAHDEBF
Subject 5	AFHGBECD	ECDFBHG*	B-A Round—angular H-E Shaped—straight tail A-D Can't give it a name E-B Linear—triangular	DBCEGHFA
Subject 6	AGBFHDEC	BDCGEHF*	D-F Moving—massive base F-E Indented on the left—not indented E-F Unstable—stable	HDCBFGEA

* For three subjects the first dimension accommodated all triads containing A. In these cases the second dimension does not contain A.

Fig. 3

The thing that strikes us immediately about them is that they are almost impossible to name. We can describe them laboriously, of course, e.g. saying of the first one that the nose becomes less tiny and grows stronger, that the tail becomes less sharp and better formed, that the body becomes less slanted and more vertical, and of the second that there is a change from hanging to standing, that the form becomes less long and thin, and fuller-bodied, that there is a progressive change from instability to stability. But although we can see what is happening, as though a piece of rubber were being deformed, step by step, our words are hardly adequate.

The fact that we have only imprecise words for what we see here, is most important, since it brings us back to the point first mentioned in the introduction: that people

* Among the very few triads not accommodated by the first two dimensions, one or two are even inconsistent—Subject 1's AFB, for instance. Probably these minor vagaries are the result of the subject's indecision already discussed, and would be smoothed out if the subject were given still longer opportunity to reach consistent choices.

do not see only in terms of the words they have for a situation. Compare the introspective dimensions offered by the subjects, with the similarity dimensions we have extracted to describe what they actually did (Fig. 3). As we see, the degree to which subjects were verbally aware of what they were doing with their eyes and hands, is very limited. Moreover, while, as we have just seen, one subject's behaviour was

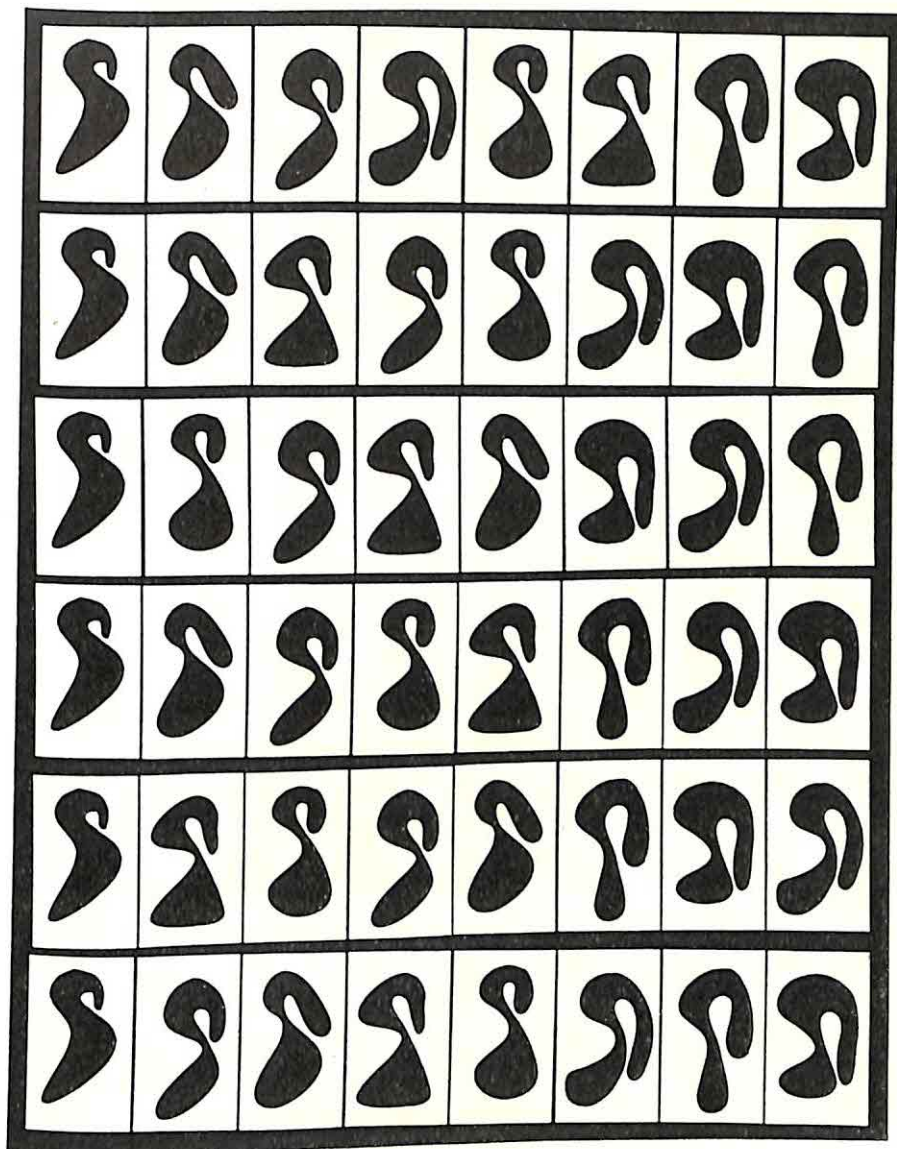


Fig. 4a. Dimension 1 for the six subjects.

much the same as another's, there is no such agreement from one subject to the next in the introspective dimensions offered. Indeed, these introspective dimensions seem to be irrelevant to the behaviour they were supposed to describe. They are connected, very certainly, with the subject's education, but unconnected, to any significant degree, with his visual behaviour.

Now for the central issue: to discover whether there is a connexion between a subject's liking for the forms and the ways in which he looks at them.

Preference orders generated by paired comparisons (the subjects' liking for the forms) are given in Fig. 3. Similarity dimensions (the subjects' ways of looking at the forms) are also given in Fig. 3. We shall test the possibility of a connexion in two different ways.

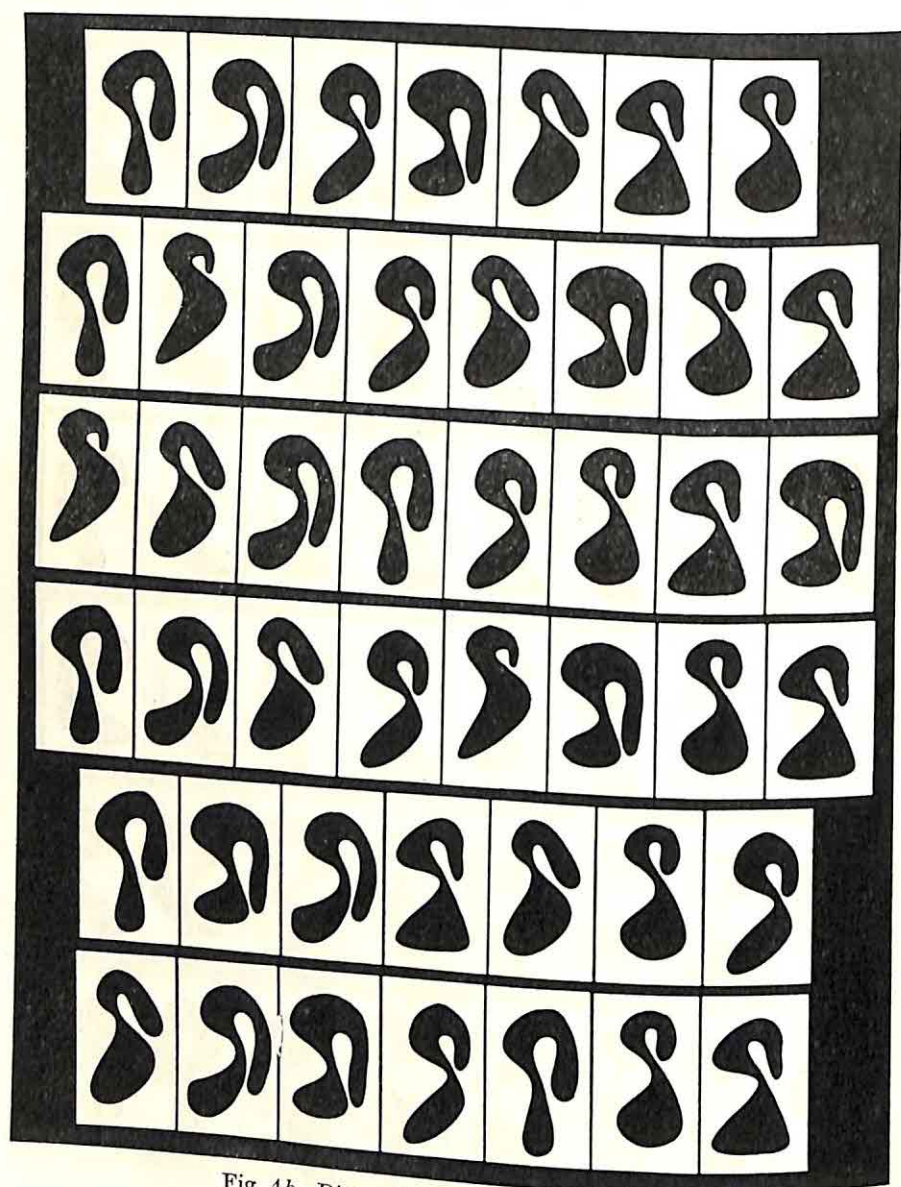


Fig. 4b. Dimension 2 for the six subjects.

(1) In the first test we shall examine each subject's preference order for its dependence on the similarity dimensions which describe that subject's behaviour. To do this we calculate the correlation between preference order and similarity dimensions, using the rank correlation coefficient τ (Kendall, 1948). These coefficients are presented in Fig. 6.

As we see, they are evenly distributed about zero, and only one (marked with a star) is above the 5% level of significance. Since among twelve coefficients we should

expect about one to reach this level, the test, points to no connexion whatever between the preference orders and the similarity dimensions.

(2) The second test is a more stringent one. For each subject there are certain triads which satisfy betweenness on both his similarity dimensions. These are listed in Fig. 7. Consider any such triad XOY (attached to a specific subject). O lies between X and Y on both his similarity dimensions. That is to say, in whichever of the two ways (or in whatever combination of them) he looks at the three forms, O will be intermediate between X and Y.

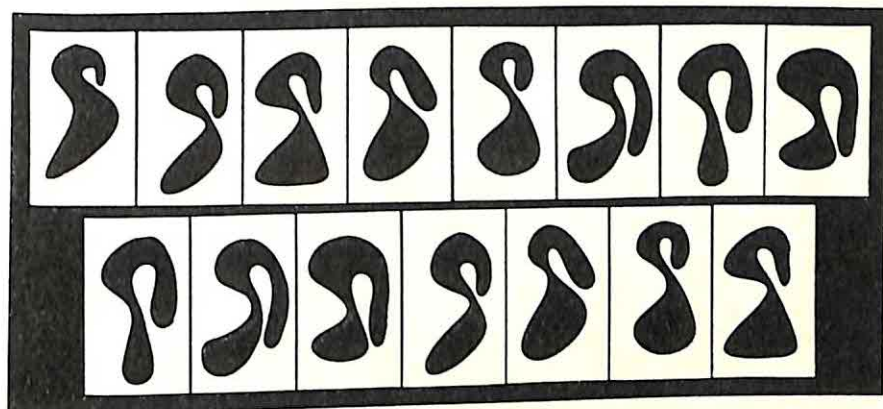


Fig. 5. Dimension 1 (modal); dimension 2 (modal).

Subject	Correlation of preference order with dimension 1	Correlation of preference order with dimension 2
Subject 1	0.00	+0.52
Subject 2	+0.29	-0.07
Subject 3	+0.14	0.00
Subject 4	0.00	0.00
Subject 5	-0.86*	+0.14
Subject 6	-0.36	+0.14

Fig. 6

If his preference order is positively connected with these dimensions, then, whatever the nature of the connexion, O should lie between X and Y on the preference order too.

In Fig. 7 we see how many of these triads do in fact satisfy betweenness on the preference orders: in all, ten out of fifty-nine. Now even if the triads were chosen at random we should expect a third of them to satisfy it; twenty out of fifty-nine, that is. So what do we make of our result? Certainly there is no positive connexion of the kind we were looking for. To establish that there would need to be twenty-seven out of fifty-nine for the 5% level of significance. We have a situation which is quite the reverse, since the number quoted is significantly *less* than random (chi-square indicates a 0.7% level of significance). This is a fact which demands explanation.

Subject	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	
Triads satisfying betweenness on both similarity dimensions	BDE	ABC*	ABC	ABD	CBH	BDC	
	BGE	ABF	ABD	ABE	DBG*	BDE	
	CED	ABH	ABE	AGD	DBH*	CEH	
	EFH*	ADC	AFC	AGE	EBG	CHF	
		AGC	AGC	BGC	EBH	DHF	
		AGH*	AHC	BGF	FHG*	EHF	
		BDE	ADE	BHF			
		BGD	AGF	BGH			
		BGE	AHF	CDE*			
		ECF	BDE*				
		ECH	BGH*				
		CHF	CFG*				
		EDF	CFH*				
		EDG	DGH				
		EDH	EGH				
		DGF					
		DHF					
		EGF					
		EHF					
Number of such triads	4	19	15	9	6	6	Total 59
Number of these triads satisfying betweenness on the preference dimension (marked above with an *)	1	2	4	1	3	0	10

Fig. 7

IV. DISCUSSION AND CONCLUSIONS

There are eight forms in the experiment. To account for a subject's sorting behaviour we therefore might need as many as seven dimensions—if his 'looking' were rich enough to involve seven ways at once. But we find that in every case two dimensions are enough to account for his entire behaviour. When asked to *say* how they looked at the forms, subjects all put forward more than two ways, it is true. But this belief in the subtlety of their looking must have been largely wishful thinking—for what they actually *did* could be described with only two dimensions.

There must be no misunderstanding at this point. The introspective dimensions offered by the subject are not meaningless. If he wants to, he can look at the forms in these ways—indeed, he *can* look at the forms in any way he pleases, for whatever verbal dimensions we make up he can set himself to place the forms along them (Osgood *et al.*, 1957). But what is essential is what he actually *did*. And what he did can be described with only two dimensions. If some third 'say' had been important to him, this fact would have been reflected in his sorting behaviour; and we should need a third dimension to account for it. (Very likely, if there had been more than eight forms, we should have needed more than two dimensions to describe his behaviour.) What is clear at any rate is that looking, as a process of categorization, is simpler than we think.

The dimensions that we extract are not verbal ones, but visual. They have no names and, directly, we cannot discuss them. But we can see them with our own eyes: we can see that something is changing from one end of the dimensions to the other, but we are hard put to it to give the 'something' a name. Often, in fact, there

are no ready words to describe the ways in which we look at forms. What about the subject's introspective dimensions—the ways in which he *thought* he had been looking at the forms? It turns out that although there is some agreement between these introspective (verbal) dimensions, and the visual ones we have extracted from his sorting behaviour, this agreement is very limited, a phenomenon which reminds us of the little girl and the Venetian paintings, but far more conclusive. We do not see only according to the way we think. On the contrary, we do not have words for what we do with our eyes.

We cannot describe our visual behaviour introspectively; and it seems that it makes good sense to refer to visual concepts which are non-verbal. What is more, while the subjects all have the same visual concepts, they have widely different verbal ones. They differ in their descriptions of their own seeing behaviour, even though the behaviour itself is much the same for all of them, a fact which does not speak well for the view that seeing is based on a learned net of language. On the contrary, the most plausible explanation is that we all share the same sort of perceptual apparatus, but have all been brought up differently, and have different words for similar visual phenomena. Our verbal concepts are largely personal—our visual ones are not.

It is one of the gifts of the great critic that, by coining words or putting old words to new uses, he can name dimensions we all use with our eyes but which we have not yet been able to name for ourselves (Wölfflin, 1915). And the painter's gift is greater still, for he makes us see (use) dimensions that we not only have no word for, but do not even know with our eyes.

Is there a connexion between a subject's liking for the forms, and the ways in which he looks at them? There is no significant correlation between the preference orders and the similarity dimensions. Nor do triads which satisfy betweenness on both similarity dimensions satisfy it on the preference dimensions, as they should if there were a positive connexion of any sort between them. We must say, therefore, that there is no positive connexion. However, the number of such triads satisfying betweenness on the preference dimension is smaller than we should expect from chance. It looks, indeed, as though a triad satisfying betweenness on both similarity dimensions will tend not to satisfy it on the preference dimension—a sort of negative interaction.

Perhaps we can illuminate this by restating it. Forms that lie at the centre of both similarity dimensions tend to lie toward the ends of the preference dimension. When a form lies at the centre of both our similarity dimensions we either like it or dislike it, but are not indifferent to it. We feel strongly about such a form. And conversely, forms at the ends of similarity dimensions tend to be neither especially beautiful, nor especially not so.

It is difficult to make much of this information. Art critics feel something similar, perhaps, when they say that a form containing divergent elements is particularly good if these elements are successfully unified, but if the contrast between them is not resolved, the form is particularly bad. Forms which lie at the centre of similarity dimensions may be said to contain diverging elements (namely, the two ends of the dimension). And the contrast between these elements will be resolved or unresolved, the form good or bad, but not indifferent.

So it appears that there is a connexion after all. Forms which lie at the centre of

both similarity dimensions tend to induce strong feelings in the subject. So much we can say. What we do not know, however, is whether these strong feelings will be favourable or unfavourable; whether the forms will be liked or disliked. While the similarities seen do restrict the preference order, they do not determine it. The preference dimension is dependent on the similarity dimensions—but the dependence is incomplete. It might, perhaps, be better called a linkage—comparable to the linkage between hue and brightness (yellows tend to be brighter than blues or reds, though hue and brightness are independent otherwise). It has been shown that attributes may often be linked to one another in very subtle ways, without being what we should normally call dependent (Stevens, 1934). In this case the weakness of the linkage is twofold:

(1) The connexion is not complete and would better be called a tendency.

(2) There is an ambiguity at the crucial point of the connexion, for we cannot predict whether the strong feeling induced will be liking or dislike.

It is well known that it is difficult to explain the beauty of forms, and that those explanations which *are* offered, are at best partial ones. Perhaps the difference we see between good forms and bad is essentially irreducible to any other differences we see, just as the difference we see between two hues of equal brightness and saturation is irreducible. For what does a successful explanation of visual quality depend upon? That our aesthetic discrimination can be made dependent on the other discriminations of which we are capable. That there is some unambiguous mapping from the other qualities our eyes allow us to pick out, onto the aesthetic one.

Yet the experiment suggests that just this is not the case. The mapping is many-many, or, as we put it, no more than a linkage. The aesthetic explanation it allows can be no more powerful therefore, than explanations of the difference between blue and yellow couched in terms of brightness could be.

This hypothesis puts no restriction on explanations which are possible in terms of experimental psychology or physiology. But it does suggest that explanations in terms of other visible qualities are essentially restricted by the weakness of the linkage.

The author wishes to thank Jerome Bruner most sincerely for his encouragement and guidance.

REFERENCES

- BROWN, R. (1956). Language and categories, appendix to *A Study of Thinking*, by J. S. Bruner, J. J. Goodnow and G. A. Austin. New York: Wiley.
- COOMBS, C. (1958). An application of a non-metric model for multidimensional analysis of similarities. *Psychol. Rep.* 4, 511–18.
- HAYS, W. (1959). The analysis of non-metric multidimensional similarities data, mimeographed. Dept. of Psychology, University of Michigan, awaiting publication 1959.
- HENDERSON, R., KATES, W. & ROHWER, W. (1959). Paintings and perception; a cognitive approach to differential perception and preference for conventional and art objects. Unpublished honours thesis, Harvard University.
- KELLY, G. A. (1955). *The Psychology of Personal Constructs*. New York: Norton.
- KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Griffin.
- OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.

- STEVENS, S. S. (1934). The attributes of tones. *Proc. Nat. Acad. Sci., Wash.*, 20 (7), 457-9.
- TUCKER, W. T. (1955). Experiments in aesthetic communications. Unpublished doctoral thesis. University of Illinois.
- WHORF, B. L. (1941). The relation of habitual thought and behaviour to language. From *Language, Culture, and Personality, Essays in Memory of Edward Sapir*, edited by Leslie Spier. Menasha, Wisconsin, Sapir Memorial Publication Fund.
- WÖLFFLIN, H. (1915). *Principles of Art History*. English translation by M. D. Hottinger. London: Bell, 1932.

(Manuscript received 10 December 1959)

PUBLICATIONS RECENTLY RECEIVED

L'Année Psychologique, 1958, No. 1, 1958, No. 2, 1959, No. 1, 1959, No. 2. Paris: Presses Universitaires de France. Pp. 319, 273, 344, 338. 1500 old francs, 1600 old francs, 18 new francs, 18 new francs.

These latest volumes of the *Année Psychologique* have the same tripartite division into Original Reports and Notes, Critical Reviews, and Bibliographical Analysis as earlier volumes.

The first section contains write-ups of experiments, all of which were apparently carried out in Paris. They bear witness to the vigorous, workmanlike and far-ranging interest with which experimental psychology is pursued at the University of Paris and with respect to which that university has at present few rivals on the European continent. Successive issues of the *Année Psychologique* show increasing contact with psychological developments in other, particularly English-speaking, countries. There is also less evidence of the indifference to systematic theory that has previously marked the Paris school. Four articles report work done by Flament within the framework of contemporary small-group theory, with special attention to its mathematical aspects. Two articles, one by Piéron and one by Galifret, are devoted to Stevens's recent reformulation of the psychophysical law, although they express reservations about it. Other articles have grown out of psychophysiological work on arousal, learning theory and Piaget's studies of perceptual illusions.

The articles under the heading of Critical Reviews demand comparison with those published in the *Psychological Bulletin*. This comparison redounds to the favour of the latter, especially when one considers breadth and thoroughness of coverage. Nevertheless, their smaller load of detail and their relative concentration on exposition rather than on evaluation may give the French review articles some advantages for certain purposes. Particular mention must be made of a highly useful summary by Heissler of recent Russian experiments on the orientation reaction, one of the most fruitful areas of concentration of contemporary Soviet psychology and physiology.

It is, however, particularly in its Bibliographical Analyses, consisting of critical abstracts of selected journal articles and books, that the *Année Psychologique* plays a role that is not quite duplicated anywhere else in psychological periodical literature. In this role, it can perform important services not only for psychologists who wish to keep informed about French psychology, but also for those who wish to keep abreast of psychology in general. These Bibliographical Analyses are not so tantalizingly laconic as many of the references in the *Psychological Abstracts* and the *Annual Review of Psychology*. They differ from the usual book-review sections of journals in having material on cognate matters grouped together and in giving priority to adequately summarizing the content. Through its geographical location and its evident wealth of linguistic accomplishment, the Parisian team of abstractors is able to give due representation to material from a wide variety of countries. Its reporting of judiciously selected Russian books and articles is especially valuable. In fact, the English-speaking psychologist who can read French but not Russian may well find the *Année Psychologique* unsurpassed as a means of keeping informed about psychology in the U.S.S.R. He is almost bound, in addition, to come across references to English-language writings in his field of specialization that he would otherwise have missed.

D. E. BERLYNE

Psychology: A Study of a Science. Study I, Vol. 3. *Formulations of the Person and the Social Context*. Edited by SIGMUND KOCH. New York: McGraw-Hill. 1959. Pp. x + 837. 97s.

Prof. Koch is the Director of a large-scale project sponsored by the American Psychological Association. The aim is to produce 'a report of inquiries into the status and tendency of psychological science'. It will appear in seven volumes of which the first three, comprising Study I, have come out; the next three, comprising Study II and dealing with the relation of psychology to other disciplines, are on the way, and in the final volume Prof. Koch will review the situation.

This volume is the third of Study I, of which the other two dealt with the theories of perception and learning. The contributors to all volumes included in Study I were issued with 'suggested discussion topics'. They were, one might say, put on the spot. They were asked to say how their theories grew in their minds, what evidence there is for them, what research they have inspired, how they stand with respect to quantification, and how they fit into the independent-intervening-dependent variable pattern. As one might expect, not all contributors follow the scheme—or, rather, in all except two cases they follow it each in his own way.

In one of the two 'deviant' cases, the paper by Franz J. Kallmann on the 'Psychogenetic Studies of Twins', the scheme is inappropriate. We are provided with an instructive account of the methodology of twin research and of the major findings so far as they are relevant to psychology. Prof. Asch, the other 'deviant', abandons the scheme altogether and discourses pleasantly enough about the data of social psychology and its relation to non-social psychology.

In other papers the scheme has produced extremely interesting results. The authors are not concerned with providing empirical information; they are concerned with theoretical concepts and their justification. It is impossible to give a summary of each contribution; all that can be done is to indicate what the papers are about.

Henry A. Murray is autobiographical. He describes the stages he passed through in his 'Preparation for the Scaffold of a Comprehensive System' of personality. A man of wide sympathies and scientific honesty, he refused all along the line to close his eyes to features of the personality which did not fit into a simple 'scientific' framework. If, say, creativity—in which he insists—does not fit in, then the scaffolding must be raised to another story.

The next paper on 'The structure of Psychoanalytic Theory', by David Rapaport is, in my view, the most interesting of them all. He 'only touches on the theory of symptom, psychosexual development, therapy (e.g. transference and resistance) and concentrates heavily on what Freud called metapsychology'. He calls his paper a 'systemizing attempt'; it is an attempt to show that psychoanalytic theory can be formulated systematically and that it hangs together as a psychological theory. He makes a very good case, and it is a relief to read a really serious treatment of psychoanalytic theory, equally free from piety as from contempt.

Carl R. Rogers describes how he came to 'A Theory of Therapy, Personality, and Interpersonal Relationships, as Developed in the Client-centred Framework', from his own experiences of counselling. It is, one may say, a pragmatic statement: this is what I found in practice, and I am forced to a certain theory about the 'fully functioning person' from my experience with incompletely functioning people—this is the mood of the paper, and what he has to say about interpersonal relations will be of practical value to all engaged in 'counselling' of one kind or another. It is the one paper which is of practical as well as theoretical interest.

Raymond B. Cattell is concerned with 'Personality Theory Growing from Multivariate Quantitative Research'. It is, in fact, about factor analysis as applied to personality, and particularly with the rationale of the method.

Theodore M. Newcomb considers the postulate that 'there are lawful interdependences among certain classes of beliefs and attitudes held by the same individual'. These interdependences are related to the strain produced by the perception of discrepancies between one's own attitudes towards an object and those of other people with whom one identifies oneself. The results provide his title: 'Individual Systems of Orientation'.

David Katz and Ezra Stotland, in their 'Preliminary Statement of a Theory of Attitude Structure and Change' discuss the same subject from a rather more general point of view.

The contributors were instructed, as has been said, to attend to the application of the independent-intervening-dependent variable pattern. In this matter the status of the intervening variable and its relation to empirical data has presented a problem to the scientific methodologist. Paul F. Lazarsfeld is concerned with just this problem, what he calls 'Latent Structure Analysis'. His paper is highly technical and seeks to provide a scheme for assessing the effectiveness of an intervening variable—trait or attitude—as an 'explanation' of the grouping of empirical material.

With Herbert A. Thelen we move into a very different field of research: the study of emotional undercurrents in small groups, in particular the relation between the work aspect of group behaviour and these undercurrents. This problem had been already in his mind for some time when Bion's papers appeared in *Human Relations*, and his contribution, called 'Work-emotionality Theory of the Group as Organism' is largely based on Bion's hypotheses. He argues that the emotional interchanges between members of groups is so closely interlocked that the only profitable way of dealing with them is to treat the group as though it were an organism. This is

no 'mere' theorising; we are given evidence; methods of prediction are described and experimental results provided.

The last paper is, perhaps, something of a surprise to those who have not followed Talcott Parsons and his school in their 'Approach to Psychological Theory in Terms of the Theory of Action'. Those who do not care for Parsons's highly abstract approach to problems of behaviour will not care for his contribution. Those who are fascinated by his extraordinary ingenuity will find it fascinating. Behaviour for Parsons belongs to an organism, a personality or psychological system, and, usually, to a social system at the same time, in which it is controlled by a cultural system. The social system and the cultural system have so far been the main concern of the Parsonians. They now turn to personality. The first point is that all four systems are *systems*, behaviours related in different ways. The second point is that, according to Parsons, the same concepts apply to all action systems. Now in his view systems of action are 'constrained by four fundamental sets of exigencies or 'functional problems''. These are: goal attainment, adaptation, integration, and pattern maintenance. These, therefore, must apply to the psychological system. We seek goals, we adapt ourselves by acquiring general skills and wealth—the general facility for acquisition—we must integrate all our needs, and maintain some coherent pattern of life. People may differ in the prominence of any of these aims and so we get a typology. But all functions must be present and so there will be an input-output interchange between the adaptive subsystem and the goal-oriented subsystem and so on with the other subsystems. It is all very neat and tidy, but the trouble, as usual, is that there are far too few empirical examples.

I have dwelt a little longer on this final paper because it is the oddest, and even so I have given only a few indications of its content. All the other contributions, with the exception of the one provided by Prof. Asch, which is on the slender side, are of the highest quality and distinction.

W. J. H. SPROTT

The Western Mind in Transition. By FRANZ ALEXANDER, M.D. New York: Random House. 1960. Pp. xvi + 300. \$5.00.

It is the privilege and often the habit of our budding septuagenarians to salute the apogee of their life's work either by writing their autobiography or by communicating to the world at large their views on its current condition and prospects. Sometimes, as in the present instance, these techniques are combined. *The Western Mind in Transition* is at the same time a professional autobiography, a thumb-nail sketch of the development of psychoanalytical thought and an apothegmatic survey of current cultural trends, combined with a few carefully safeguarded speculations about the future.

Let it be said at once that few are as qualified to exercise the privilege of ex-cathedra commentary as Franz Alexander. Born and bred in the stimulating atmosphere of the University of Budapest, nourished in the traditions of philosophic humanism, his earlier ventures in scientific method led him finally to the field of psychoanalysis where he found his spiritual home. Trained in metaphysical speculation, he took to psychoanalytical metapsychology like a duck to water. When the present reviewer first met Alexander in his early thirties, he was already the white-headed boy of the Berlin Psychoanalytical Society and for the next forty years played a prominent part in the teaching and practice of psychoanalysis, at first in Berlin and latterly in the United States.

At the same time, it is to be observed that the evaluation of social and cultural trends is a risky profession. The interrelation between subjective thought and feeling and objective (social) experience is so close, particularly in the unconscious regions of the mind, that what starts as a personal reaction often ends by masquerading as a social generalization. How often do our publicists maintain that the community is 'gravely concerned' about this and that, when in fact only the author is concerned, leaving the inhabitants of Main Street to go about their daily business of living quite unperturbed by the misdirected diagnosis.

Dr Alexander is of course well aware of this danger and confines himself mainly to such generalizations as, he feels, can be supported by psychoanalytical experience. Following the lead of Erik Erikson, his main concern is with the 'loss of ego-identity' which follows an unsuccessful struggle for equilibrium between a loosely organized self and a changing environment. Following

the history of the individual's integrative processes from childhood onwards, he distinguishes between the 'inner-directed' person who after maturation has a surplus of energy to spend freely and the 'other-directed' person whose self has never crystallized and who, taking his cues from outside, is constantly engaged in making rapid adaptations to rapidly changing social conditions. 'Other-directed' types have, he maintains, little energy left for individual creativity and real human contact. This contrast has been rendered more acute by the modern emergence of communal and conformist man living in an era of automation and mass production without stable principles of conduct and subject always to the pressures both of his fellow (communal) men, and of increasingly totalitarian organizations of society. The Western mind, in short, is in a state of transition which threatens the creative individuality of the unintegrated.

This theme the author pursues through a number of chapters dealing with such matters as psychoanalysis, the role of the scientist in society, revolutions in science and art, existentialism, organization and the role of play in creative work, with a penultimate section on what he calls the 'retirement neurosis'. Dr Alexander is, however, no pessimist. He believes that when man has adapted to this stage of transition, he will again proceed to build on more abundant economic foundations and, guided by his interests in art, in science for its own sake, in education and in research will utilize his capacity for play to find fresh expressions for his creative mind. So after some tense moments we may all breathe freely.

Needless to add, Dr Alexander believes that during this process of reorientation, the insights of psychoanalysis will exert a favourable influence, provided always psychoanalysts refrain from becoming methodolaters and pay more attention to the immediate dynamic needs and problems of the individual. This adjuration, it may be noted, is becoming quite a theme-song amongst neo-Freudians, existential psychiatrists and those eclectics who find their consulting rooms overflowing with disgruntled and inadequate 'characters'. Indeed it is more than probable that a change in clientele, more notable in the U.S.A. than elsewhere, is responsible for the 'self-other' slanting of much modern psychotherapy and perhaps also for the sense of transition which has so impressed Dr Alexander.

All in all this book makes good bedside reading for psychologists and those general readers who like to ponder the nature and effects of change without being called on for too strenuous mental effort. Dr Alexander is sufficiently persuasive to prevent 'inner-directed' readers being alienated by his exercise in 'other-direction'. Even if they should disagree with his diagnosis and maintain that during historical times man has continued to be in a state of transition without manifesting any very profound change in his inner nature and mechanisms, they must concede that the author has stated his case with vigour, zest, clarity and reminiscential bonhomie.

EDWARD GLOVER

Mechanisation of Thought Processes. Proceedings of a Symposium held at the National Physical Laboratory in November 1958. Vols. 1 and 2. London: H.M. Stationery Office. 1959. Pp. x+980. £2. 10s. 0d. (two volumes).

It is pointed out in the preface that 'this symposium was held to bring together scientists studying artificial thinking, character and pattern recognition, learning, mechanical language translation, biology, automatic programming, industrial planning and clerical mechanisation'. Rather fewer than one-half of the thirty-two papers presented are of direct interest to psychology. Even so, the symposium shows that interest in psychological problems associated with cybernetics is not flagging.

Session 1, devoted to General Principles, includes papers by three pioneers in cybernetics and related fields, D. M. MacKay, W. R. Ashby and A. M. Uttley, and two mathematicians from the Massachusetts Institute of Technology, M. L. Minsky and J. McCarthy. MacKay's paper is an extension of his several earlier suggestions of 'a self-organizing information-system'. Ashby's interesting paper considers the mechanism of habituation. He sets out to show that 'when any system is subjected to a regularly repeated stimulus or disturbance, the successive responses, if they change in size, do not in general tend to become larger or smaller with equal probability: there is a fundamental bias in favour of the smaller'. Uttley is concerned with the elucidation of the mathematical principles which 'may underlie the organization of nervous systems'; his particular

subject is 'conditional probability computing in a nervous system'. Minsky and McCarthy are concerned with the programming of computers to make them exhibit different forms of seemingly intelligent behaviour. In fact, the whole of Session 2, comprising seven further papers, is also devoted to problems of programming. Papers in this section, however, deal with specific technical problems rather than with general principles. The first four are contributed by mathematicians, including A. P. Ershov of the University of Moscow. Two further papers in this section are concerned with mechanical language translation, and the last is a very brief report by Ershov on 'Automatic translation in the U.S.S.R.'

There are two groups of papers in Session 3: three on speech recognition and three on learning in machines. Speech recognition is here considered mainly from the point of view of physiology and phonetics, and the problems discussed would be familiar only to those relatively few psychologists who are actually engaged in research in this field. It is now well known that, adopting one view of learning or another, it is possible to think of machines that will learn. F. Resenblatt of Cornell Aeronautical Laboratory, O. G. Selfridge of Lincoln Laboratories, and A. M. Andrew of the National Physical Laboratory describe such machines. Whether our understanding of learning processes is advanced by this approach is still an open question. The seven papers in Session 4A are said to be concerned with 'Implications for Biology'. In fact, four of the papers are presented by experimental psychologists. Of the other three, one, entitled 'Medical Diagnosis and Cybernetics', is by F. Paycha of Paris; another paper by H. B. Barlow suggests how large information inputs are handled by sensory systems; in the third paper the great veteran, W. S. McCulloch returns to the subject of nervous nets. The psychologists are keen to be taken as biological scientists, trying to discern the mechanics of nervous and behavioural functioning. N. S. Sutherland's paper, 'Stimulus Analysing Mechanisms' and R. L. Gregory's, 'Models and the Localization of Function in the Central Nervous System' are concerned with nervous functioning; the paper by A. J. Watson, 'Some Questions Concerning the Explanation of Learning in Animals', and J. Brown's paper, 'Information, Redundancy and Decay of the Memory Trace', are very much in the tradition of theorizing in experimental psychology.

There are seven more papers in Session 4B, 'Implications for Industry'. Most of them are concerned with such topics as automation in the legal world, mechanization of administration, etc. Rather different is G. Pask's paper on 'Physical Analogues to the Growth of a Concept'; perhaps this paper did not naturally fit into any other part of the symposium. The two volumes contain, in addition, discussions on all the papers, reports of five lecture-demonstrations, an attendance list, and an index to authors and contributors to discussions.

W. SLUCKIN

Culture and Mental Health. Edited by MARVIN K. OPLER. New York and London: The Macmillan Company, New York. 1960. Pp. xxi + 533. 61s.

This is an omnibus volume in which Dr Opler has brought together some twenty-three papers concerned with the variable effects of culture and cultural stress on mental health. Of the studies sixteen have not been reported before. The contributors are drawn from many disciplines; they include anthropologists, psychiatrists, sociologists, psychologists and psychoanalysts.

Reports from a refreshing wide variety of different cultures find a place in this volume. There are good sections on the North and South American Indian and on some Asian communities, while Africa and Micronesia are less well represented. There are some interesting accounts of American ethnic groups in which the mental health of such minority groups as Negroes, Spanish-Americans, Jews, Irish and Italians is compared with that of suitably selected samples of the native American population. A final section on 'World Perspective' includes a thumb-nail review of recent research sponsored by W.H.O. and U.N.E.S.C.O.

As one might expect, a wide range of research techniques have been employed and while this adds interest it does so at the expense of unity. It is on this score that the book was most disappointing. One missed the sense of an underlying ground-plan to the book. Grouping the papers into rough geographical areas does little to enable the reader to understand the basic problems of transcultural psychiatry. In his Preface the editor discusses his problems in organizing the book. In deciding not to group methodologically similar studies together and accompany them with a discussion of the techniques thus illustrated, the editor has perhaps lost the opportunity of giving us a text-book which both the University teacher and the practising clinician would

undoubtedly have found most useful. As it is, the book, well written and interesting though it is, may very well prove to be too unstructured, too long and too expensive to find its way into many bookshelves.

J. DREWERY

Delinquency and Parental Pathology. By ROBERT G. ANDRY. London: Methuen. 1960. Pp. xv + 173. 21s.

This is a comparison of child-parent relationships in a group of eighty delinquents with those among a similar number of non-delinquent controls. Its inspiration was the maternal deprivation hypothesis of Dr John Bowlby; its aim, to show that this hypothesis is one-sided, and that paternal deprivation is equally important in the making of delinquents.

The method was to examine, in interview, by means of a schedule of standard questions, the child's perception of his parents' current relationships with him. Thus questions were asked about such topics as which parent gives more love, or shares leisure-time occupations with the child. Physical separation from the parents was also investigated as well as aspects of the child's own behaviour which give some measure of the degree of social adjustment he has been able to achieve—the latter being seen as consequences of the pattern of child-parent relationships experienced by him. Delinquency is the main factor studied here. From the larger sample thus studied, a sub-sample of thirty delinquents and thirty controls was selected for further investigation. In these cases, both parents were interviewed, partly as a check on the accuracy of the child's responses, and partly to obtain information about his early training experiences.

The results obtained are of great interest, and they certainly do realize the author's subsidiary objective: 'to determine what differences exist between delinquents and non-delinquents in regard to the adequacy with which each parent plays his/her role'—at least as that role is reported on by the child himself. Clear differences between delinquents and controls do occur. In particular the father's role is reported by the children as less adequately played among the delinquents than the non-delinquents. A number of other very interesting facts emerge—one being that 68 % of the controls had stolen at one time or another. This calls attention once more to the inadequacy of conviction data as a measure of the amount of delinquency, or as a guide to the individuals or groups who are delinquent. It also emphasizes the pitfalls which lie in wait in this field for the unwary framer of research hypotheses. It is, for example, hardly satisfactory to call the two groups studied in the present research, 'delinquents' and 'non-delinquents'.

The further aim of the book, to challenge the maternal deprivation theory, is not, in this reviewer's estimation, nearly so well achieved as that of discriminating between convicted delinquents and those who have not been so convicted. For one thing, it is hazardous to claim that because a thirteen-year-old child feels his father less affectionate at the moment, that this means anything at all about the quality of father-child relationships when the child was between 6 and 18 months of age. Yet it is the latter age with which the maternal deprivation hypothesis is primarily concerned.

There is also, however, a methodological objection of some importance. To achieve the rigour and repeatability aimed at in this study, a highly standardized procedure has to be adopted. There is no scope for clinical exploration, which is so necessary if the real feelings of the child on these strongly felt and highly subjective matters are to be understood. Clinical experience suggests that, in, for example, a matter which is as vital to the child as his mother's attitude to him, he may even deceive himself about what he has perceived in his family situation. It is no answer to say, as the author of this book seems to be saying at one point, that if standardized methods cannot test a theory so much the worse for the theory. That is too much like 'throwing the baby out with the bath water'.

HOWARD JONES

The Education of Slow Learning Children. By A. E. TANSLEY and R. GULLIFORD. London: Routledge and Kegan Paul. 1960. Pp. viii + 255. 28s.

Since the war there has been an increasing interest in the educational needs of handicapped children. Numerically the educationally subnormal are the largest group and wartime separations focused attention on the problems of poor educational attainments and illiteracy. Since then considerable expansion has taken place in the provision of special schools and special classes

for educationally subnormal pupils. This book is primarily addressed to teachers of such children. The authors stress that there is no clear line of demarcation between the better pupils in an E.S.N. school and the most backward ones in ordinary schools. Just as the intellectual calibre of children in Grammar schools varies very considerably according to the number of places available in local selective schools, so one finds considerable variations in the types of children admitted to E.S.N. schools depending on the extent and variety of special educational facilities provided in any one area. Moreover, since the term 'educational subnormality' is a broad one, the authors wisely address themselves to teachers of all types of slow learning pupils.

After a brief historical account of the principles underlying special educational treatment, the intellectual, emotional, physical and social development of slow learning children is discussed. Then methods of teaching the basic subjects, of fostering language development and creative ability are outlined with many practical suggestions and examples. Though research findings are referred to only briefly, there are quite full bibliographies at the end of each chapter. This book achieves a successful synthesis of educational psychology and educational practice. Well written and at a moderate level of difficulty, it should be of interest not only to teachers but also to those responsible for their training.

M. L. KELLMER PRINGLE

Der Mongolismus. By KARL KÖNIG. Stuttgart: Hippokrates Verlag. 1959. Pp. 278. DM 29.

Of all the ante-natal developmental disorders, mongolism is perhaps the most important. Its appearance, or perhaps rather its discovery, as a specific manifestation of mental deficiency, is as recent as a hundred years ago. Since then its incidence has increased so that now it takes first place among all severe types of mental deficiency. Mongolism also constitutes a serious psychological and educational problem as such children usually come from normal families. The book by Dr König is therefore of interest particularly since he has, for many years, been a pioneer in the treatment and education of mongols.

Part I consists of a review of the literature. It is argued that mongolism should not be regarded as a developmental disorder but rather as an early, incomplete form of human development, like a butterfly arrested at the larval stage. Both mind and body have remained 'unfinished, too simple and too coarse to serve the subtle and complex needs of everyday life'.

In Part II the author presents his analysis of the birthdays of some 2700 mongoloid children. While the labour involved in this analysis is impressive, this reviewer feels somewhat sceptical about its value, particularly regarding the link between dates of birth and phases of the moon. It is perhaps not surprising that no comparable analysis exists for the births of normal children so that Dr König cannot determine whether the patterns he discovered are general or apply only to mongols.

Part III outlines developmental stages in physical, emotional and intellectual growth. The data are based on Dr König's examination of 120 mongols and he gives comparative figures for the same number of children suffering from other mental handicaps, such as epilepsy, schizophrenia and post-encephalitic conditions. He claims that the recent advances by Ford and others in understanding the nature of mongolism confirms his own contention that the condition is not so much a defect as the occurrence of a special type of human being, a throw-back to a precursor of homo sapiens. Therefore he regards mongolism as primarily a sociological rather than a medical problem.

The last part of the book deals with the treatment and education of mongols. There Dr König argues from his work in residential schools, all of which are based on the principles of Rudolph Steiner. In his view most mongols can benefit from education by the age of eight to nine years. However, the teaching methods he advocates make it clear that training is given rather than education, since it is conceded that most of the children will never acquire real understanding and insight.

It is very difficult to assess the value of this book to the psychologist and educationist. On the one hand it brings together material and ideas from biological, genetic, medical, and sociological fields. On the other hand, it is an unusual, and at times curious, mixture of scientific inquiry, theosophy and mysticism. There is no subject index, only an index of authors.

M. L. KELLMER PRINGLE

Research Methods in Social Relations. By CLAIRE SELLTIZ, MARIE JAHODA, MORTON DEUTSCH and STUART COOK. New York: Holt; London: Methuen. 1960. Pp. xvi + 622. 42s.

This is a revised one-volume edition of a work first published in 1951, which gained a well-deserved reputation. The general framework of the original first volume has been retained, but its content has been considerably expanded. Some of the less technical parts of the earlier second volume have been included, supplemented by more recent approaches to methodological problems and brief accounts of new techniques. Illustrations are no longer confined to the field of inter-group relations.

Although the declared purpose of the revision was to make the book more suitable for undergraduate courses, there is none of the spoon-feeding one commonly associates with this. The first few chapters on formulation of research problems and research design are in fact quite rigorous and make rather stiff reading; there are, for instance, excellent discussions of such topics as the rationale of randomization and matching, and types of 'before-after' designs. The lucid manner in which the concepts of validity and reliability are dealt with in the broader context of problems of measurement is also noteworthy.

The series of chapters on direct and indirect methods of data collection and scaling are comprehensive in their coverage, but several sections suffer perhaps from excessive condensation; thus some readers will regret that the detailed treatment of sociometry in the first edition has been replaced by a few sketchy pages. However, some compensation is afforded by the inclusion of material on the Q-sort and the semantic differential. Among the most valuable parts of the work is that concerned with problems of analysis and interpretation: it contains an unusually good coverage of such questions as categorization of data, and potential errors in inferring causal relations. In discussing tabulation, the authors point out that for some purposes hand sorting of small cards may be quite effective, but the absence of any reference to Cope-Chat cards makes one wonder whether these may be missing from the armoury of American researchers. Finally, mention may be made of the appendices on budgeting for a study, sampling, questionnaire construction and interviewing. All these are thoroughly practical, as shown, for instance, by the following statement, whose depressing truth can hardly be repeated often enough: 'It appears to be an almost universal rule that every operation takes longer than one would anticipate if everything went smoothly'.

In general, the authors have succeeded in conveying not only the formal aspects and mechanics of research, but also something of its flavour. The book is so densely packed that one can hardly advise beginners to tackle it without a good deal of guidance; on the other hand, there are probably few seasoned researchers who cannot learn anything from it.

G. JAHODA

Essentials of Psychological Testing. By LEE J. CRONBACH. New York: Harper and Brothers. 1960 (2nd edition). Pp. xxi + 560. \$7.00.

The author's remark in the preface to the book that 'there has been substantial change from the first edition' is very much an understatement. Every section has been reviewed and brought up to date, and a wealth of new material, drawing on the 'notable advances in testing and test theory' of the past decade, has been added. Every possible aspect of testing is considered, sometimes at great length, making this one of the most comprehensive books on the subject at present available. And to add that the present volume is better than its predecessor is indeed to praise it.

The first part of the book deals with 'basic concepts'—purposes and types of tests, administration, scoring, validity and reliability—and since it is primarily intended as a text-book for students, the elementary statistics necessary for understanding these concepts are graphically and painlessly introduced. Every point is demonstrated with reference to existing tests—unfortunately nearly all American—so that throughout the reader can form an opinion as to the practical efficiency of the tests mentioned and at the same time appreciate the theoretical framework on which the science of mental measurement is based.

The second and third parts of the book deal with testing ability and performance—individual and group tests of ability, differential aptitude tests, personal selection, tests of attainment and interest, personality measurement and clinical approaches. Again, throughout this latter part of the book the reader is introduced to specific techniques by way of the practical considerations of testing requirements. For example, an excellent description of factor analysis is sandwiched between a discussion of general mental ability and differential aptitude testing. While this logical positioning is to be commended it is a little confusing if the book is to be used for reference purposes, and it is probably in this way that it will be used. Not many people will wish to read the 650 pages from cover to cover, although all those directly concerned with mental measurement will find this a rewarding task.

English students of psychology may find the book a little on the 'heavy' side, not because the syntax is difficult—on the contrary, the book is extremely readable—but because it deals so exhaustively with every aspect of the subject. This is not meant as a criticism, rather it is a comment on what English students are expected to know about testing. It is to be hoped, perhaps, that teachers in this country as well as the taught will refresh their ideas from this book—critics of testing not excepted.

D. A. PIDGEON

Scoring Human Motives: A Manual. By J. DOLLARD and F. AULD. Yale University Press; London: Oxford University Press. 1960. Pp. x + 452. 76s.

This book is not, as its title might suggest, concerned with ordinary everyday behaviour but with psychotherapy. It presents in great detail a scoring system for categorizing the verbal and behavioural transactions of the psychotherapeutic interview—at least so far as these survive sound-recording and subsequent typewritten transcription. Since the authors tell us that 'the time is almost at hand when every scientific student of behaviour will be to some degree a Freudian', it is no surprise to find that the scoring categories lean heavily on psychoanalytic ideas.

About a dozen categories are offered for the patient's words and behaviour, and half as many for those of the therapist. Each of the scoring symbols employed is carefully defined and exemplified and its underlying theory discussed at length. 'A', for example, indicates 'anxiety or apprehension'; 'L', 'love, liking, approaching'; and 'W', 'weeping'. The symbols are combined in various ways ('L-A' indicates 'A loving emotion to which anxiety is attached') and a distinction is made between conscious processes (indicated by capital letters) and unconscious processes (small letters). A good deal of interpretation is called for from the scorer, as, for example, when allocating the symbol 's-a:r' (which scores a unit of behaviour in which 'an unconscious anxiety attached to unconscious sexual wishes is reduced'). Considerable attention is devoted to 'unitizing' or breaking up the typewritten record into scorable units. Criteria for assessing the reliability of the procedure are discussed at length and impressive inter-observer agreement is reported. The authors consider the validity of the procedure by comparing interviews of psychotics and neurotics, by studying covariation between categories which psychoanalytic or learning theory suggest might be linked, and by studying changes in the frequency of certain categories over the course of therapy.

This is a bold and energetic attempt to deal with a notoriously difficult field and the failure of the authors to deal convincingly with unconscious processes is hardly unexpected. The book is much too long and expensive, but it might provoke some useful argument about problems of categorization—and perhaps some useful research.

JOHN LIGGETT

Die Sprache der Zeichnung: Der Wartegg-Biedma-Test—Eine Abwandlung des Wartegg-Tests. By CARLOS J. BIEDMA and PEDRO G. D'ALFONSO. Bern and Stuttgart: Huber. 1959. Pp. 110. Sw. Fr./DM. 24.80.

This modification of the original Wartegg Drawing Test may be described as the reverse of Gottschalt's hidden figures. Instead of the isolation of objects masked in a larger configuration, subjects are encouraged to integrate a given symbol which may be a dot, an arc, parallel lines

or a square into a larger drawing of their own choice. The completed drawings are analysed in three ways:

(1) presentation—involving the objective drawing characteristics and assessment of static vs. dynamic content;

(2) integration—assessed in terms of graphic 'archetypes' basic drawing symbols which the authors claim to be present in 'all graphic, plastic and architectural configurations irrespective of time or place having an identical psychological significance'; and

(3) projection—drawing content seen in relation to interpreted themes and attitudes.

The analysis is mainly scored according to presence or absence of objective criteria, and the interrelationship of the three parameters is described in terms of an arbitrarily derived set of arithmetical indices.

The authors present their projective technique with a great deal of faith in its sensitivity as a tool for the assessment of personality and character traits. Beyond mentioning the use of Rotter's Incomplete Sentences as a validating device, no evidence of external and diagnostic validity is given nor considered necessary, within the ideographic frame of reference of the technique.

While the present attempt to objectify the notoriously difficult procedure of scoring individual drawings is of considerable interest, a major validation programme is clearly desirable before the technique can be used confidently in diagnosis and prediction. The book includes many scored examples, test blanks and material for scoring.

VERNON HAMILTON

10,000 Careers. By ROBERT L. THORNDIKE and ELIZABETH HAGEN. New York: John Wiley and Sons, Inc. 1959. Pp. x + 346. 68s.

This book is probably as important as the Kelly and Fiske study and will probably be misinterpreted just as often. Thorndike and Hagen report results of following up 10,000 men who had passed through USAAF selection procedure twelve years before. The predictor variables were the scores from twenty tests and biographical data on a Blank; the criteria were obtained from a postal questionnaire which revealed present job, income, level of responsibility, self ratings of success and job satisfaction, vertical and lateral mobility and length of time in job.

A wealth of statistical skill was brought to bear on this wealth of material. It is impossible in a short review to do justice to the results but here are a few. Occupational groups (for example, engineers, business executives) showed differences, not so much in the level of mean test scores and other variables as in the unevenness of the profiles. Items on the Biographical Blank, particularly educational and social background, differentiated between occupational groups at least as well as test scores did. Tests did not predict success within occupations.

Thorndike and Hagan emphasize again and again the limitations of this study, the most important of which are these: the sample was a highly selected one representing 'roughly the top half of a high school graduating group'; replies came back from only 70 % of the sample and there was evidence that those who did not reply were less successful in their civilian careers; the occupational groups were very heterogeneous; only cognitive test results were available and there were no measurements of temperament, personality and interests; the criteria were obtained from the men's own reports on themselves, employers not being consulted.

However, even with the limitations—and no study has ever lacked them—the results of this investigation must be considered more chastening than encouraging. Occupational psychologists should study the book carefully.

DENIS MCMAHON

The Lore and Language of Schoolchildren. By IONA and PETER OPIE. Oxford: Clarendon Press. 1959. Pp. xix + 417. 35s.

When it appeared (a year ago) the Opies' book was so widely and appreciatively reviewed in the national press that a more technical review seems hardly to be called for. Nevertheless, this copy has been so thoroughly read by this reviewer's children, aged from seven upwards, that it seems only fair to record their appreciation, and one's own. There are, in addition, many points of psychological or at any rate sociological interest which might repay more systematic

study; how one would set about it is another matter—although certainly using an approach very different from the author's. Among the questions they leave unanswered are such things as the precise nature of the grape-vine whereby nearly identical versions of a popular item turn up in remote parts of the world almost more quickly than modern communications could convey them; just what conception children have of popular or notorious figures (Kruger, Kitchener, Felix the Cat, Deanna Durbin) no longer in the public eye but living on in children's lore; to what extent the efficacy of 'magical' practices is really believed in, and so on. Comparable phenomena are of course observable in adult life, but it seems likely that the implications may be different.

The rhymes, customs, etc., collected by the Opies come mostly from Great Britain, excluding Ireland, but items are noted from many other parts of the world. Naturally enough a relatively small number of 'contributing schools' have been responsible for major portions of the data, but the frontispiece map shows that substantial contributions have been obtained from nearly all parts of England, Wales and Scotland. One's main complaint would be that neither the charming photograph on the dust jacket nor any others appear in the book itself.

B. SEMEONOFF

Psychological Analysis of Interest. By HELENA SLONIEWSKA. The Poznan Society of Friends of Science. Department of Philology and Philosophy. Publications of the Section of Philosophy. Vol. 10, No. 1. Poznan. 1959. Pp. 110. In Polish, with a summary in English. 22 Zloty.

This is a study of interests of educated adults by a lecturer in psychology in the University of Wroclaw. The subjects studied were university students and members of the staff. Each of the subjects was given a catalogue test, i.e. a list of titles of books with the instruction to choose ten such as he or she would like to read. The test was followed by an interview on the modes in which the subject's interests as indicated by his choice of titles were reflected in his thought and behaviour, such as reading, conversations and cultural pursuits. The tests and interviews took place in 1950.

Information elicited in the interview suggests a definition of interest as readiness to occupy oneself intellectually with an object. Interest as readiness is aroused by a wide variety of stimuli. In the situation investigated interest is stirred even by titles of books bearing only indirectly on the object in which the person is interested. Thus we come to the distinction of direction of interest indicated in this case by the choice of titles—and its real object.

A further distinction is to be drawn between the object of an interest and its background constituted by conscious causal factors and motives often of personal nature. Confusion of object and background has led previous students to asserting of interest what more properly should be asserted of its background, so that it was impossible for them to agree how interest was to be defined.

Studying the background of various interests of the same person we discover ties binding the distinctive interests into one mental whole varying in structure from individual to individual and closely connected with personality.

Interests can be divided into personal and objective; the latter into practical and theoretical interests. Likewise, the background of an interest can be classified as mainly personal or objective, practical or theoretical.

It does not come as a surprise to learn that interest in sex, art, ethics and religion is personal, while interest in philosophy, science and technology is mainly objective. But it does come as a mild shock to find that the background of interest in psychology which has proved to be very much alive also in persons who do not engage in psychology professionally—is personal, mainly personal unhappiness. It is significant that in the background of interest in philosophy, ethics and science there was in the minds of these people a need for some consistent outlook on life and code of conduct. The year was 1950. In this respect the book is an important, maybe unique document on psychology of social and political upheaval.

A. BARDECKI

- Descartes: Discourse on Method and other Writings.* Translated with an Introduction by ARTHUR WOLLASTON. Harmondsworth: Penguin. 1960. Pp. 192. 3s. 6d.
- Psychology and Religion.* By C. G. JUNG. Newhaven: Yale University Press; London: Oxford University Press. 1960. Pp. 131. 7s. 6d.
- Becoming.* By GORDON W. ALLPORT. Newhaven: Yale University Press; London: Oxford University Press. 1960. Pp. ix + 106. 7s. 6d.
- The Origins of Love and Hate.* By IAN D. SUTTIE. Harmondsworth: Penguin. 1960. Pp. xvi + 222. 3s. 6d.
- The Wish to fall Ill: A Study of Psychology and Medicine.* By KARIN STEPHEN. Cambridge University Press. 1960. Pp. vii + 238. 10s. 6d.
- Our Adult World and its Roots in Infancy.* By MELANIE KLEIN. London: Tavistock. 1960. Pp. 15. 3s. 6d. (Tavistock Pamphlet No. 2).

Originally grouped together as a selection of recent paper-backs, all of them reprints, the above also show a certain degree of uniformity of subject-matter, within a broad general field embracing religion, psychoanalysis and theories of personality.

Doubtfully eligible for inclusion, perhaps, is the late Arthur Wollaston's new translation of Descartes. The 'other writings' are the *Meditations* and the 'Letter Preface' to the *Principles of Philosophy*. The translator has added a 25-page Introduction and a brief 'Note on Cartesian Terminology', and has prefaced each section of the work with an expository summary.

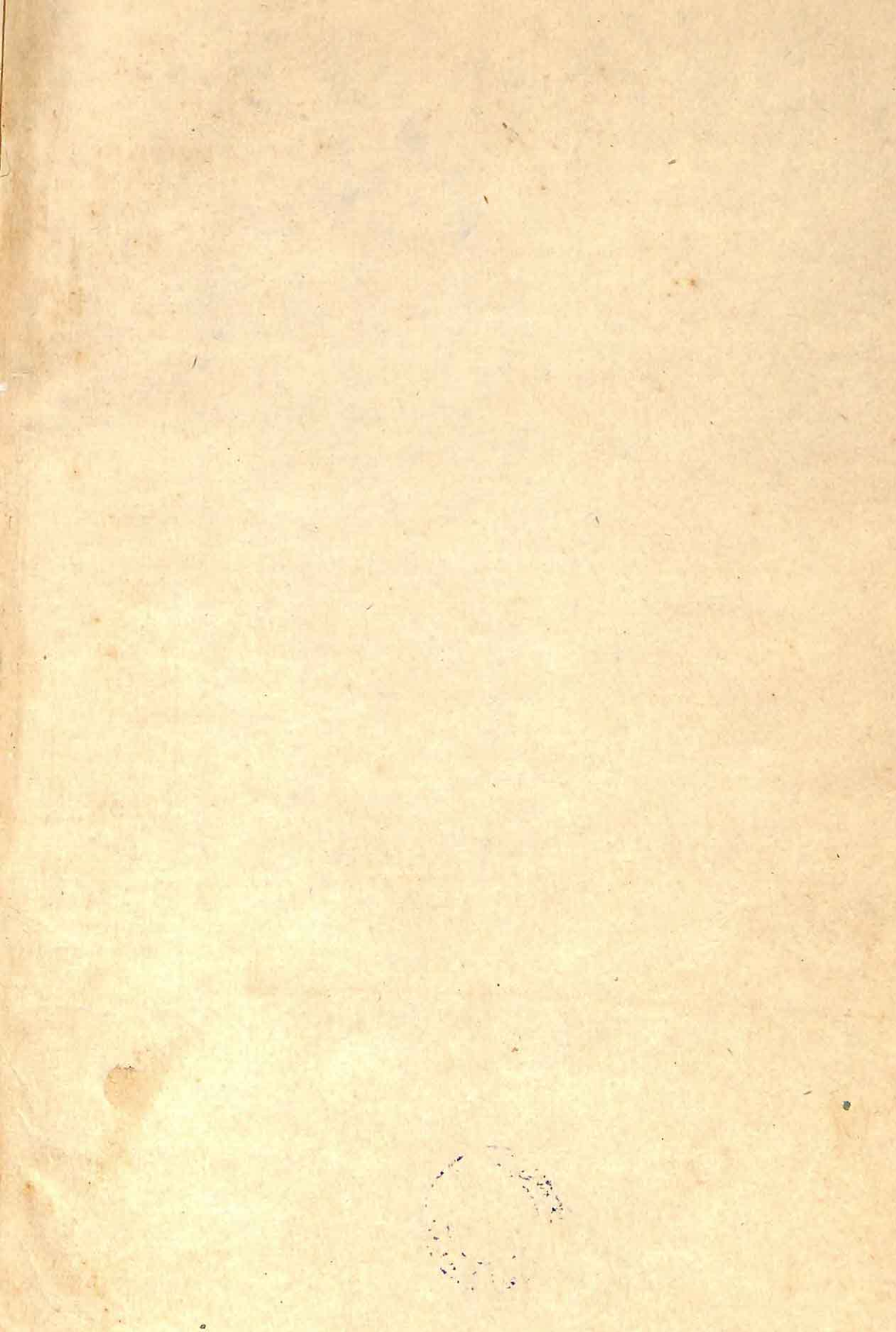
The two 'Yale Paperbounds' present the text of the lectures given at Yale under the auspices of the Terry Lecture Foundation in 1939 and 1954 respectively. To quote Prof. Allport: 'The Terry lecturer is given the assignment of assimilating and interpreting his discipline as it relates to human welfare and to religion broadly conceived.' A link with Descartes can be found in the words of a reviewer of the original publication of Allport's work, who saw in it 'a theory of the psyche inclusive enough to encompass both the hopes of the eighteenth century and the fears of the twentieth'.

Suttie's book and Karin Stephen's, both of which appeared originally in the mid-thirties, may be regarded as complementary to one another, since the former offers a severe but constructive criticism of Freudian psychology, and the latter (which originally had the two parts of its title in the reverse order) is an account of the psychoanalytic view of the structure and dynamics of the human mind, with particular reference to psychogenic illness.

An Editorial Note introduces the Tavistock Pamphlet as follows: 'This paper, originally delivered by Melanie Klein on 11 May 1959 before the members of the Departments of Social Anthropology and Social Studies in the University of Manchester, is a brief but comprehensive statement of her findings and theories in psychoanalysis in a form specially prepared for a wider social-science audience.' The present publication is a straight reprint from an article in *Human Relations*. This seems an excellent idea, and one which, if extended to important papers in other journals, might go far towards solving the problems of authors who have to deal with embarrassing and—possibly—not always justifiable requests for offprints.

B. SEMEONOFF





✓



1066

